DISSERTATION

QUANTILE REGRESSION MODELS

OF ANIMAL HABITAT RELATIONSHIPS

Submitted by

Brian S. Cade

Graduate Degree Program in Ecology

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Spring 2003

COLORADO STATE UNIVERSITY

December 16, 2002

WE HEREBY RECOMMEND THAT THE DISSERTATION PREPARED UNDER OUR SUPERVISION BY BRIAN SCOTT CADE ENTITLED QUANTILE REGRESSION MODELS OF ANIMAL HABITAT RELATIONSHIPS BE ACCEPTED AS FULFILLING IN PART REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY.

Committee on Graduate Work

____

____

____

____
Advisor

____
Department Head

ABSTRACT OF DISSERTATION

QUANTILE REGRESSION MODELS OF
ANIMAL HABITAT RELATIONSHIPS

Typically, all factors that limit an organism are not measured and included in statistical models used to investigate relationships with their environment. If important unmeasured variables interact multiplicatively with the measured variables, the statistical models often will have heterogeneous response distributions with unequal variances. Quantile regression is an approach for estimating the conditional quantiles of a response variable distribution in the linear model, providing a more complete view of possible causal relationships between variables in ecological processes. Chapter 1 introduces quantile regression and discusses the ordering characteristics, interval nature, sampling variation, weighting, and interpretation of estimates for homogeneous and heterogeneous regression models. Chapter 2 evaluates performance of quantile rankscore tests used for hypothesis testing and constructing confidence intervals for linear quantile regression estimates ($0 \leq \tau \leq 1$). A permutation F test maintained better Type I errors than the Chi-square T test for models with smaller n, greater number of parameters p, and more extreme quantiles τ. Both versions of the test required weighting to maintain correct Type I errors when there was heterogeneity under the alternative model. An example application related trout densities to stream channel width:depth. Chapter 3 evaluates a drop in dispersion, F-ratio like permutation test for hypothesis testing and constructing confidence intervals for linear quantile regression estimates ($0 \leq \tau \leq 1$). Chapter 4 simulates from a large (N = 10,000) finite population representing grid areas on a landscape to demonstrate various forms of hidden bias that might occur when the effect of a measured habitat variable on some animal was confounded with the effect of another unmeasured variable (spatially and not spatially structured). Depending on whether interactions of the measured habitat and unmeasured variable were negative (interference interactions) or positive (facilitation interactions), either upper (τ >0.5) or lower (τ < 0.5) quantile regression parameters were less biased than mean rate parameters. Sampling (n = 20 - 300) simulations demonstrated that confidence intervals constructed by inverting rankscore tests provided valid coverage of these biased parameters. Quantile regression was used to estimate effects of physical habitat resources on a bivalve mussel (*Macomona liliana*) in a New Zealand harbor by modeling the spatial trend surface as a cubic polynomial of location coordinates.

Brian S. Cade
Graduate Degree Program
In Ecology
Colorado State University
Fort Collins, CO 80523
Spring 2003

ACKNOWLEDGMENTS

I would like to express my appreciation for the support and help of my committee members, Drs. Barry R. Noon (advisor), Paul W. Mielke, Jr., Kurt D. Fausch, and Curtis H. Flather. My research was funded and supported by the Fort Collins Science Center, U. S. Geological Survey. I want to thank James E. Roelle, Richard L. Schroeder, and James W. Terrell for their helpful critique and review of my work. Jon D. Richards provided wonderful computer programming support. I also wish to express my gratitude to Roger Koenker for tutoring me in the essentials of quantile regression. I appreciate the support and love of my family, Amy, Jessica, and Sam.

TABLE OF CONTENTS

ABSTRACT OF DISSERTATION..iii

ACKNOWLEDGMENTS v

Chapter 1. A GENTLE INTRODUCTION TO QUANTILE REGRESSION FOR ECOLOGISTS

Abstract...1

Introduction...2

Quantiles and ordering in the linear model...6

Homogeneous and heterogeneous models...8

Estimates are for interval of quantiles...13

Sampling variation differs across quantiles...15

Second order properties of the estimates are useful.. 17

Discussion... 18

Literature cited..20

Chapter 2. RANKSCORE AND PERMUTATION TESTING ALTERNATIVES FOR REGRESSION QUANTILE ESTIMATES

Abstract...26

Introduction..27

Quantile regression model..........31
Rankscore test statistics..........33

Simulation experiment..........38

Example application..........58

Discussion..........65

Literature cited..........68

Appendix 2..........75

Chapter 3. A DROP IN DISPERSION PERMUTATION TEST FOR REGRESSION QUANTILE ESTIMATES

Abstract..........88

Introduction..........89

Quantile regression model..........91

Test statistics..........92

Simulation experiment..........94

Example application..........109

Discussion..........112

Literature cited..........114

Appendix 3..........119

Chapter 4. ESTIMATING EFFECTS OF LIMITING HABITAT RESOURCES: HIDDEN BIAS AND SPATIAL STRUCTURE

Abstract..........124

Introduction..........125

Quantile regression models with unmeasured variables...129

Patterns of effects due to confounding with unmeasured processes.............................132

Performance of regression quantiles tests for models with hidden bias......................143

Example application..152

Discussion...172

Literature cited..178

Appendix 4..184

Chapter 1

A Gentle Introduction to Quantile Regression for Ecologists

Abstract: Typically, all factors that limit an organism are not measured and included in statistical models used to investigate relationships with their environment. If important unmeasured variables interact multiplicatively with the measured variables, the statistical models often will have heterogeneous response distributions with unequal variances. As a consequence, there may be no or weak predictive relationship between the mean of the response variable (*y*) distribution and the measured environmental factors. Yet, there may be stronger, useful predictive relationships with other parts of the response variable distribution. Quantile regression is an approach for estimating the conditional quantiles of a response variable distribution in the linear model, providing a more complete view of possible causal relationships between variables in ecological processes. This introduction relates quantile regression to estimates of prediction intervals in parametric error distribution models (e.g., least squares regression) and discusses the ordering characteristics, interval nature, sampling variation, weighting, and interpretation of the estimates for homogeneous and heterogeneous regression models. The motivation is to address the large variation often found in relationships between ecological variables and the presumed causal factors that is not attributed to random sampling variation. These models are useful when the response variable is affected by more than one factor, factors vary in their effect on the response, not all

factors are measured, and the multiple limiting factors interact.

1. Introduction

Regression is a common statistical method employed by scientists to investigate relationships between variables, where a response variable y is some function of predictor variables X, $y = f(X)$. Most regression applications in the ecological sciences, whether linear or nonlinear in the parameters or nonparametric, focus on estimating rates of change associated with the mean of the response variable distribution as some function of a set of predictor variables, i.e., the function is defined for the expected value of y conditional on X, $E(y|X)$. Mosteller and Tukey (1977) noted that it was possible to fit regression curves to other parts of the distribution of the response variable, but that this was not commonly done and, thus, most regression analyses gave an incomplete picture of the relationships between variables. Heterogeneous variances are pervasive in regression models used to estimate relationships between variables in ecology. An exclusive focus on effects (regression slope coefficients) associated with changes in the means may under estimate, over estimate, or fail to distinguish real nonzero changes in heterogeneous response variable distributions (Terrell et al. 1996, Cade et al. 1999).

Regression quantiles were developed by econometricians in the 1970's (Koenker and Bassett 1978) as a straight forward, semiparametric extension of the linear model to estimate rates of change in all parts of the distribution of the response variable. They are semiparametric in the sense that no parametric distributional form (e.g., normal, Poisson, negative binomial, etc.) is assumed for the error distribution. Recent literature

(Cade et al. 1999, Koenker and Machado 1999, Koenker and Hallock 2002) denotes the quantiles by the Greek τ, where $\tau \in [0, 1]$, although this notation is by no means universal. The conditional quantiles denoted by $Q_y(\tau|X)$ are the inverse of the conditional cumulative distribution function of the response variable, $F_y^{-1}(\tau|X)$. For example, for $\tau = 0.90$, $Q_y(0.90|X)$ is the 90th percentile of the distribution of y conditional on the values of X, i.e., 90% of the values of y are less than or equal to the specified function of X. Note, that for symmetric distributions the 0.50 quantile (or median) is equal to the mean μ. Here I consider functions of X that are linear in the parameters, e.g., $\beta_0(\tau)X_0 + \beta_1(\tau)X_1 + \beta_2(\tau)X_2 +, ..., + \beta_p(\tau)X_p$, where the ($\tau$) notation indicates that the parameters are for a specified τ quantile. The parameters vary due to effects of the τth quantile of the unknown error distribution ε. Quantile regression provides a very flexible method of modeling the rates of change in the response variable at multiple points of the distribution for both homogeneous and heterogeneous error models, providing a much more complete picture of the relationships between variables (Koenker and Machado 1999).

In the 1-sample setting with no predictor variables, estimating quantiles is usually thought of as a process of ordering the sample data. The beauty of the extension to the regression model was recognizing that quantiles could be estimated by an optimization function minimizing a sum of weighted absolute deviations, where the weights are functions of τ (Koenker and Machado 1999, Koenker and Hallock 2002). Currently, the statistical theory and computational routines for estimating and making inferences on regression quantiles are best developed for the linear model

(Gutenbrunner et al. 1993, Koenker 1994, Koenker and Machado 1999) but also are available for parametric nonlinear (Welsh et al. 1994, Koenker and Park 1996) and nonparametric, nonlinear smoothers (Koenker et al. 1994, Yu and Jones 1998). Improved methods of testing hypotheses and inverting hypothesis tests for constructing confidence intervals on parameters of linear regression quantile models are the topics of Chapters 2 and 3.

There have been a variety of applications of quantile regression in ecology and biology, including studies of animal habitat relationships (Terrell et al. 1996, Haire et al. 2000, Eastwood et al. 2001, Dunham et al. 2002), prey and predator size relationships (Scharf et al. 1998), body size of deep-sea gastropods and dissolved oxygen concentration (McClain and Rex 2001), vegetation changes associated with agricultural conservation practices (Allen et al. 2001), variation in nuclear DNA of plants across environmental gradients (Knight and Ackerly 2002), Mediterranean fruit fly survival (Koenker and Geling 2001), running speed and body mass of terrestrial mammals (Koenker et al. 1994), global temperature change over the last century (Koenker and Schorfheide 1994), and plant self-thinning (Cade and Guo 2000). Many applications have used regression quantiles as a method of estimating functional rates of change along or near the upper boundary of the conditional distribution of responses because of issues raised by Kaiser et al. (1994), Terrell et al. (1996), Thomson et al. (1996), Cade et al. (1999), and Huston (2002). These authors suggested that if ecological limiting factors act as constraints on organisms, then the estimated effects for the measured factors were not well represented by changes in the means of response variable distributions when there were many other unmeasured factors that were potentially limiting. The response of the organism cannot change by more than some upper limit set by the measured factors but may change by less when other unmeasured factors are limiting. This analytical problem is closely related to the more general statistical issue of hidden bias in observational studies due to confounding with unmeasured variables (Rosenbaum 1995, 1999). The multiplicative interactions among measured and unmeasured ecological factors that contribute to this pattern are explored in more detail relative to regression quantile estimates and inferences in Chapter 4.

Although many of the initial ecological applications of quantile regression focused on estimating a subset of the upper regression quantiles (e.g., $\tau > 0.90$) to identify effects of limiting factors, it is possible to obtain estimates across the entire interval of quantiles ($\tau \in [0, 1]$) as a flexible method of modeling distributional changes conditional on some set of covariates. Regression quantile estimates can help reveal effects of important variables that were not measured by providing a more complete view of heterogeneous effects in the response distribution (Chapter 4). Quantile regression models present many new possibilities for statistical analyses and interpretations of ecological data (Cade et al. 1999, Cade and Guo 2000). With those new possibilities come many new challenges related to estimation, inference, and interpretation. Here I provide an overview of several of the issues ecologists are likely to encounter when conducting and interpreting quantile regression analyses. More technical discussion is provided in the relevant literature cited.

2. Quantiles and ordering in the linear model

Regression quantile estimates are an ascending sequence of planes that are above an increasing proportion of sample observations with increasing values of the quantiles τ (Fig. 1.1A). It is this operational characteristic of regression quantiles that extends the concepts of quantiles, order statistics, and rankings to the linear model (Gutenbrunner et al. 1993, Koenker and Machado 1999, Koenker and Hallock 2002) The proportion of observations less than or equal to a given regression quantile estimate, e.g., the 90th percentile given by $Q_y(0.90|X)$ in Figure 1.1A, will not in general be exactly equal to τ. The simplex linear programming solution minimizing the sum of weighted absolute deviations ensures that any regression quantile estimate will fit through at least $p + 1$ of the n sample observations for a model with $p + 1$ predictor variables X. This results in a set of inequalities defining a range for the proportion of observations less than or equal to any selected quantile τ given n and p (Cade et al. 1999, Koenker and Machado 1999).

Regression quantiles, like the usual 1-sample quantiles with no predictor variables, retain their statistical properties under any (linear or nonlinear) monotonic transformation of y as a consequence of this ordering property, i.e., they are equivariant under monotonic transformation of y (Koenker and Machado 1999). Thus it is possible to use a nonlinear transformation (e.g., logarithmic) of y to estimate linear regression quantiles and then back transform the estimates to the original scale (a nonlinear function)without any loss of information. This, of course, is not possible with means, including those from regression models.

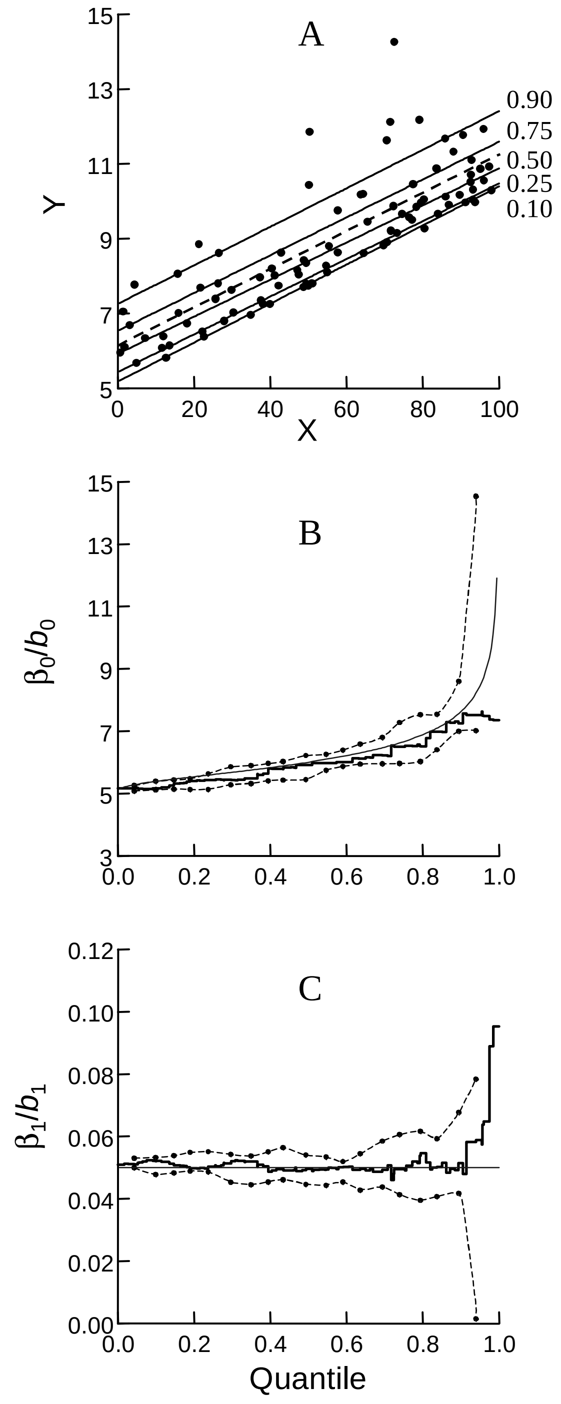


Figure 1.1 (A) is a sample (n = 90) from a homogenous error (lognormal with median = 0 and σ = 0.75) model, $y = \beta_0 + \beta_1 X_1 + \varepsilon$, β_0 = 6.0 and β_1 = 0.05 with 0.90, 0.75, 0.50, 0.25, and 0.10 regression quantile estimates (solid lines) and least squares estimate of mean function (dashed line). Sample estimates, $b_0(\tau)$ in (B) and $b_1(\tau)$ in (C), are shown as a thick solid step function. Parameters $\beta_0(\tau)$ in (B) and $\beta_1(\tau)$ in (C) are the thin solid lines. Dashed lines connect endpoints of 90% confidence intervals.

The examples presented here have been kept to simple linear regression models with just an intercept and a single predictor variable for simplicity of presentation. More complicated linear models involving a mix of categorical (indicator variables) and multiple continuous variables and their interactions are possible (Cade et al. 1999, Dunham et al. 2002). The parameter estimates in regression quantile linear models have the same interpretation as those in any other linear model. They are rates of change conditional on adjusting for the effects of the other variables in the model.

3. Homogeneous and heterogeneous models

The simplest, unconstrained form of the regression quantile estimates allows the predictor variables (X) to exert changes on the central tendency, variance, and shape of the response variable (y) distribution (Koenker and Machado 1999, Koenker and Hallock 2002). This is possible without modification of the model specified as a function of the predictor variables. When the only estimated effect is a change in central tendency (e.g., means) of the distribution of y conditional on the values of X, we have the familiar homogeneous variance regression model associated with ordinary least squares regression (Fig. 1.1A). All the regression quantile slope estimates $b_1(\tau)$ are for a common parameter and any deviation among the regression quantile estimates is simply due to sampling variation (Fig. 1.1C). An estimate of the rate of change in the means from ordinary least squares regression also is an estimate of the same parameter as for the regression quantiles. The intercept estimates $b_0(\tau)$ of the regression quantile model are for the parametric quantile, $\beta_0(\tau)$, of y when $X_1 = 0$, which differ across quantiles τ and for the mean μ (Fig. 1.1B). Intercept estimates differ across quantiles both because of sampling variation and because the parameters differ. Here the primary virtue of the regression quantile estimates of the intercept is that they are not dependent on an assumed form of the error distribution as when least squares regression is used, which assumes a normal error distribution.

The properties associated with the intercept translate to any other fixed value of $X_1, X_2, ..., X_p$ as when estimating prediction intervals for some specified value of the predictor variables (Neter et al. 1996). The interval between the 0.90 and 0.10 regression quantile estimates in Figure 1.1A at any specified value of $X = x$ is an 80% prediction interval for a single future observation. Prediction intervals for some number of future observations that assume a normal error distribution as is done in ordinary least squares regression are sensitive to departures from the distributional assumptions (Neter et al. 1996), whereas regression quantile estimates avoid this distributional assumption altogether. Given the skewness in the response distribution in Figure 1.1A it is easy to see that a symmetric prediction interval about an estimate of the mean would not have correct coverage, as would occur if we assumed a normal error distribution model. For example at $X = 70.5$ the 80% prediction interval for a single new observation is 8.43 - 10.97 based on the least squares estimate assuming a normal error distribution, whereas the interval based on the 0.90 and 0.10 regression quantile estimates is 8.85 - 10.88. Zhou and Portnoy (1996) provided an empirical evaluation of various intervals based on regression quantile estimates. Simultaneous prediction intervals for all X (tolerance bands) based on inverting quantile rankscore tests are discussed in Chapter 2 and 4.

When the predictor variables X exert both a change in means and a change in variance on the distribution of y, we have a regression model with unequal variances (a location/scale model in statistical terminology). As a consequence, changes in the quantiles of y across X cannot be the same for all quantiles τ (Fig. 1.2). The slope estimates $b_1(\tau)$ differ across quantiles both because of sampling variation and because the parameters differ since the variance in y changes as a function of X (Fig. 1.2C). Note that in this regression model with heterogeneous variances the pattern of changes in estimates $b_0(\tau)$ mirror those for $b_1(\tau)$. In this situation ordinary least squares regression is commonly modified by incorporating weights (that usually have to be estimated) that are inversely proportional to the variance function (Neter et al. 1996). Typically, the use of weighted least squares is done to improve estimates of the sampling variation for the estimated mean function, and not done specifically to estimate the different rates of change in the quantiles of the distributions of y conditional on X. However, Hubert et al. (1996) and Gerow and Bilen (1999) described applications of least squares regression where this might be done. Estimating prediction intervals for some number of future observations based on weighted least squares estimates implicitly recognize these unequal rates of change in the quantiles of y (e.g., Cunia 1987).

Generalized linear models offer alternative ways to link changes in the variances (σ^2) of y with changes in the mean (μ) based on assuming some specific distributional form in the exponential family, e.g., Poisson, negative binomial, gamma (McCullagh and Nelder 1989). But, again, the purpose usually is to provide better

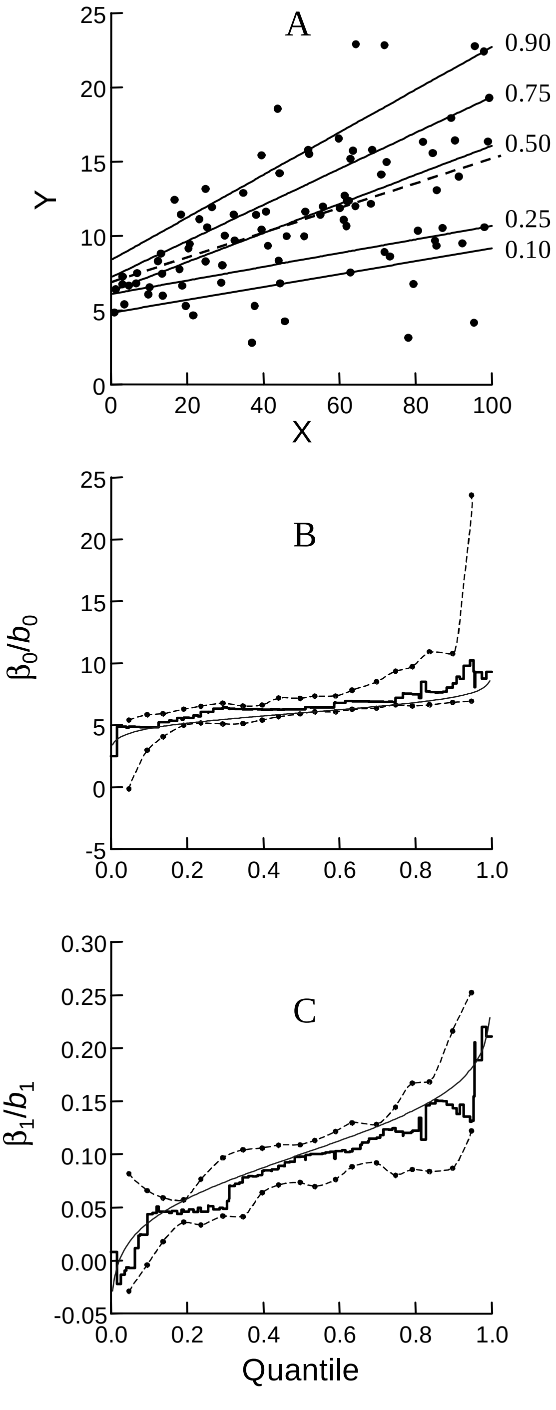


Figure 1.2 (A) is a sample (n = 90) from a heterogeneous error (normal with $\mu = 0$ and $\sigma = 1.0 + 0.05X_1$) model, $y = \beta_0 + \beta_1 X_1 + \varepsilon$, $\beta_0 = 6.0$ and $\beta_1 = 0.10$ with 0.90, 0.75, 0.50, 0.25, and 0.10 regression quantile estimates (solid lines) and least squares estimate of mean function (dashed line). Sample estimates, $b_0(\tau)$ in (B) and $b_1(\tau)$ in (C), are shown as a thick solid step function. Parameters $\beta_0(\tau)$ in (B) and $\beta_1(\tau)$ in (C) are the thin solid lines. Dashed lines connect endpoints of 90% confidence intervals.

estimates of rates of change in the mean (μ) of *y* rather than estimates in the changes in the quantiles of *y* which must occur when variances are heterogeneous. Estimating prediction intervals for generalized linear models would implicitly recognize that rates of change in the quantiles of *y* cannot be the same for all quantiles, and these interval estimates would be linked to and sensitive to violations of the assumed error distribution.

An advantage of the regression quantile approach to modeling heterogeneous variation in distributions of the responses is that no specification of how the variance changes are linked to the mean are required. Furthermore, it is possible for the predictor variables to also exert changes in the shape of the distributions (Koenker and Machado 1999, Koenker and Hallock 2002). Complicated changes in central tendency, variance, and shape of distributions are common in statistical models applied to observational data because of model misspecification. Model misspecification can occur because the appropriate functional forms are not used (e.g., linear instead of nonlinear) and because all relevant variables are not included in the model (Cade et al. 1999, Chapter 4). Failure to include all relevant variables does not necessarily occur because of scientific neglect but because of insufficient knowledge of or ability to measure all relevant processes. This should be considered the norm for observational studies in ecology as it is in many other scientific disciplines.

An example of a response distribution pattern that may involve changes in central tendency, variance, and shape is in Figure 1.3. These data from Irwin and Cook (1985) and Cook and Irwin (1985) were collected to estimate how pronghorn

(*Antilocapra americana*) densities changed with features of their habitat on winter ranges. Here shrub canopy cover was the habitat feature used as an indirect measure of the amount of winter forage available. Note that rates of change in pronghorn densities due to shrub canopy cover (b_1) were fairly constant for the lower 1/3 of the quantiles (0.25 per change in % cover), increased moderately in rate for the central 1/3 of the quantiles (0.25 to 0.50), and doubled (0.50 to 1.0) in the upper 1/3 of the quantiles (Fig.1.3C). The changes in $b_1(\tau)$ do not appear to mirror those for $b_0(\tau)$ indicating that there is more than just a change in central tendency and variance of pronghorn densities associated with changes in shrub canopy cover. Clearly, too strong a conclusion is not justified with the small sample (n =28) and large sampling variation for upper quantiles as indicated by 90% confidence intervals on the estimates. But either an ordinary least squares regression estimate (b_1 = 0.483, 90% CI = 0.31- 0.66) or more appropriate weighted least squares regression estimate would fail to recognize that pronghorn densities changed at both lower and higher rates as a function of shrub canopy cover at lower and upper quantiles of the density distribution, respectively. Here, the regression quantile estimates provide a more complete characterization of an interval of changes in pronghorn densities that were associated with changes in winter food availability as measured by shrub canopy cover. These intervals are fairly large because pronghorn densities on winter ranges are almost certainly affected by more processes than just food availability as represented by shrub canopy cover.

4. Estimates are for intervals of quantiles

Regression quantile estimates break the interval $\tau \in [0, 1]$ into a finite number of

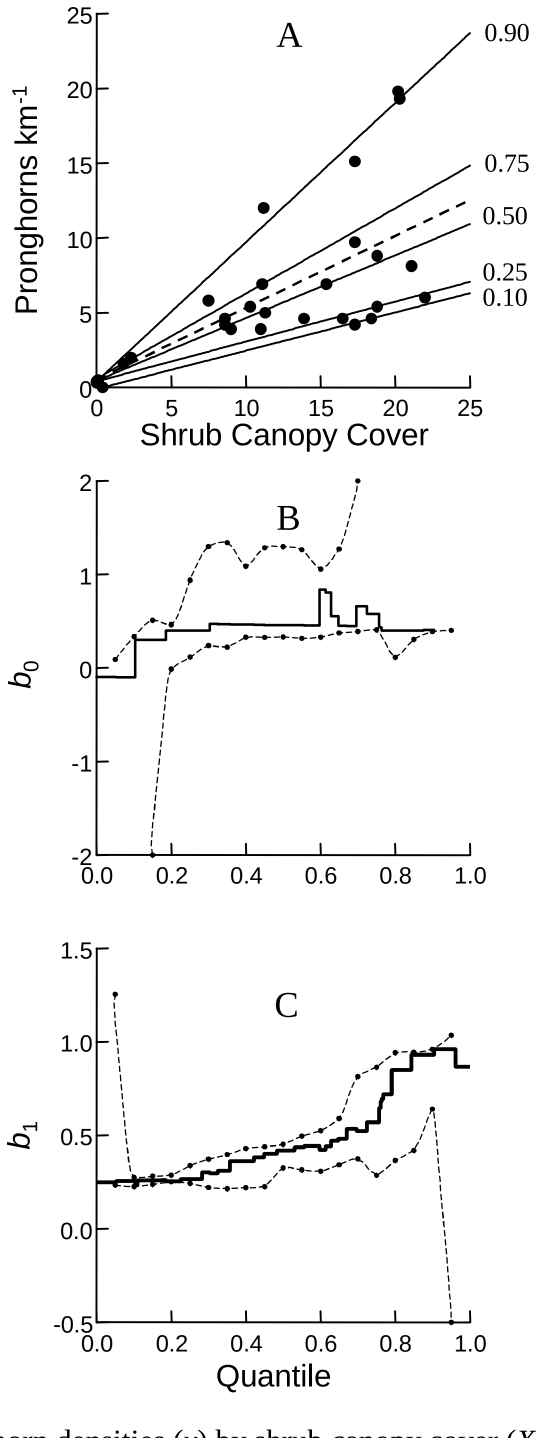


Figure 1.3. (A) is pronghorn densities (y) by shrub canopy cover (X_1) on n = 28 winter ranges (data from Cook and Irwin 1985) and 0.90, 0.75, 0.50, 0.25, and 0.10 regression quantile estimates (solid lines) and least squares regression estimate (dashed line) for the model $y = \beta_0 + \beta_1 X_1 + \varepsilon$. Sample estimates, $b_0(\tau)$ in (B) and $b_1(\tau)$ in (C), are shown as a thick solid step function. Dashed lines connect endpoints of 90% confidence intervals. Missing interval endpoints in (B) were not estimable.

smaller, unequal length intervals. Thus, while we may refer to and graph the estimated function for a selected regression quantile such as the 0.90, the estimated function actually applies to some small interval of quantiles, e.g., [0.894, 0.905] for the 0.90 regression quantile in Figure 1.1. Unlike the 1-sample quantile estimates, the [0, 1] interval of regression quantile estimates may be broken into more than *n* intervals that aren't necessarily of equal length 1/*n*. The number and length of these intervals are dependent on the sample size, number or parameters, and distribution of the response variable. Estimates plotted as step functions in Figure 1.1B and C are for 101 intervals of quantiles on the interval [0, 1] for which each has an estimate $b_0(\tau)$ and $b_1(\tau)$, corresponding to the intercept and slope. Because the estimates actually apply to a small interval of quantiles, it is appropriate to graph the estimates by intervals of quantiles as a step function (Fig. 1.1B and C). Graphing estimates as a step function by quantiles becomes essential when there are >2 predictor variables in a model. For a finite sample size *n* and $p + 1$ predictor variables $X_0, X_1, X_2, ..., X_p$ (X_0 is a column vector of 1's for an intercept), the maximum number of unique regression quantile estimates on $\tau \in [0, 1]$ is of order $n\log(n)$ (Koenker and d'Orey 1987, Portnoy 1991).

5. Sampling variation differs across quantiles

It should come as little surprise that the sampling variation can differ among quantiles τ. Generally, sampling variation will increase as the value of τ approaches 0 or 1, but the specifics are dependent on the data distribution, model fit, sample size *n*, and number of parameters *p*. Estimates further from the center of the distribution (the median or 50th percentile given by $Q_y(0.50|X)$) usually cannot be estimated as precisely. To display the sampling variation with the estimates (Fig. 1.1B and C), a confidence band across the quantiles $\tau \in [0, 1]$ was constructed by estimating the pointwise confidence interval for 19 selected quantiles $\tau \in [0.05, 0.10, ..., 0.95]$. These intervals were based on inverting a quantile rankscore test (Koenker 1994, Cade et al. 1999, Koenker and Machado 1999, Chapter 2). It is possible to compute confidence intervals for all unique intervals of quantiles but this computational effort is not usually required to obtain a useful picture of the estimates and their sampling variation. The endpoints of the confidence intervals were not connected across quantiles as a step function because they were only estimated for a subset of all possible quantiles.

Other procedures for constructing confidence intervals than the rankscore test inversion exist, including the direct order statistic approach (Zhou and Portnoy 1996, 1998), a drop in dispersion permutation test (Chapter 3), and various asymptotic methods dependent on estimating the variance/covariance matrix and the quantile density function (Koenker and Machado 1999). An advantage of the rankscore test inversion approach is that it turns the regression quantile inference problem into one solved by least squares regression for which there already exists a wealth of related theory and methods (Chapter 2).

In the example in Figure 1.1, the 90% confidence intervals for both the intercept (β_0) and slope (β_1) are narrower at lower quantiles, consistent with the fact that the data were generated from a lognomal error distribution (median = 0, $\sigma = 0.75$) which had higher probability density and, thus, less sampling variation at lower quantiles. Also note that the endpoints of the confidence intervals estimated by inverting the quantile rankscore test are not always symmetric about the estimate (Koenker 1994), which is consistent with the skewed sampling distribution of the estimates for smaller n and more extreme quantiles. The population parameters for the intercept, $\beta_0(\tau)$, and slope, $\beta_1(\tau)$, are contained within the 90% confidence intervals for most quantiles τ (Fig. 1.1B and C).

6. Second order properties of the estimates are useful

The rates of change across quantiles in the slope parameter estimates (e.g., Fig. 1.3C) can be used to provide additional information that can be incorporated into the model to provide estimates with less sampling variation. The sampling variation of a selected τ regression quantile estimate is affected by changes in the parameters in some local interval surrounding the selected quantile, say $\tau \pm h$, where h is some bandwidth (Koenker and Machado 1999). Weighted regression quantile estimates can be based on weights that are inversely proportional to the differences in estimates for some local interval of quantiles, e.g., 0.90 ± 0.06 (Koenker and Machado 1999, Chapter 4). A variety of methods have been proposed for selecting appropriate bandwidths (Koenker and Machado 1999). The difference between the local interval approach to constructing weights and estimating the variance function to construct weights as for weighted least squares regression (e.g., Neter et al. 1996:400-409) is that the former approach allows the weights to vary for different quantiles, whereas the latter approach assumes common weights for all quantiles (Chapter 4). Differential weights by quantiles are more appropriate for patterns of response similar to those in Figure 1.3 where a second order analysis suggested that rates of change in the estimates were not likely just due to changes in means and variances because the changes in $b_1(\tau)$ across quantiles did not mirror those of $b_0(\tau)$. Common weights for all quantiles are appropriate for patterns of responses similar to those in Figure 1.2 where only location and scale changes occurred as indicated by changes in $b_1(\tau)$ across quantiles that mirrored those of $b_0(\tau)$.

7. Discussion

Estimating quantiles of the response distribution in regression models is not new. This has always been required for constructing prediction and tolerance intervals for future observations, but has usually been done only in a fully parametric model where the error distribution takes some specified form. In the full parametric model the various quantiles of the response distribution are estimated by a specified multiple of the estimated standard deviation of the parametric error distribution which is then added to the estimated mean function. Vardeman (1992) stressed the importance of prediction (for some specified number of future observations) and tolerance intervals (for a proportion of the population and thus any number of future observations) in statistical applications. Much current statistical practice with linear models focuses on estimating confidence intervals on parameters. The difference between prediction/tolerance intervals and confidence intervals is that the former deal with the sampling variation of individual observations and the latter with the sampling variation of parameter estimates (which are a function of the n observations). Prediction and tolerance intervals are far more sensitive to deviations from the assumed parametric error distribution than are confidence intervals. Regression quantile estimates can be used to construct prediction and tolerance intervals without assuming some parametric error distribution and without specifying how variance heterogeneity is linked to changes in means.

The additional advantage provided by regression quantiles is to directly estimate changes in the quantiles of the distribution of responses conditional on the *p* predictor variables, i.e., $\beta_1(\tau)$, $\beta_2(\tau)$, ..., $\beta_p(\tau)$, which cannot be equal for all quantiles τ in models with heterogeneous error distributions. Differences in rates of change at different parts of the distribution are informative in a variety of ecological applications. Complicated forms of heterogeneous response distributions should be expected in observational studies where many important processes may not have been included in the candidate models. From a purely statistical standpoint, higher rates of change associated with some more extreme quantiles (e.g., $\tau > 0.90$ or $\tau < 0.10$) of the distribution may be detected as different from zero in sample estimates more often (i.e., greater power) than some central estimates such as the mean or median ($\tau = 0.50$). This can occur because greater differences between the parameter estimates and zero (no effect) can offset the greater sampling variation often associated with the more extreme quantiles. The use of regression quantile estimates in linear models with unequal variances will permit detection of effects associated with variables that might have been dismissed as statistically indistinguishable from zero based on estimates of means (Terrell et al. 1996).

The ability to statistically detect more effects with regression quantiles than conventional linear model procedures is not a panacea for investigating relationships between variables. Along with the greater ability to detect a multitude of effects comes the additional responsibility for the investigator to clearly articulate what is important to the process being studied and why. A search through all possible quantiles on a large number of models with many combinations of variables for those with strong nonzero effects is no more likely to produce useful scientific generalizations than similar unfocussed modeling efforts using conventional linear model procedures.

Finally, software is currently available to provide a variety of quantile regression analyses. Scripts and fortran programs to work with S-Plus are available from the web sites of Roger Koenker (*www.econ.uiuc.edu~roger/research/home.html*) and the Ecological Archives E080-001 (*www.esapubs.org/archive/ecol/ E080/001/ default.htm*). Add on packages for R are available from the Comprehensive R Archive Network (*lib.stat.cmu.edu/R/CRAN/*). Quantile regression estimates for linear models, quantile rankscore tests, and permutation testing variants are available in the Blossom statistical packaged available from the U. S. Geological Survey (*www.fort.usgs.gov/ products/ software/blossom.asp*). Two econometrics commercial packages that provide quantile regression are Stata and Shazam.

Acknowledgments

J. E. Roelle, R. L. Schroeder, and J. W. Terrell reviewed drafts of the manuscript.

Literature Cited

Allen, A. W., B. S. Cade, and M. W. Vandever. 2001. Effects of emergency haying on vegetative characteristics within selected conservation reserve program fields in the northern Great Plains. Journal of Soil and Water Conservation **56**:120-125.

Cade, B. S., and Q. Guo. 2000. Estimating effects of constraints on plant performance with regression quantiles. Oikos **91**: 245-254.

Cade, B. S., J. W. Terrell, and R. L. Schroeder. 1999. Estimating effects of limiting factors with regression quantiles. Ecology **80**:311-323.

Cook, J. G., and L. L. Irwin. 1985. Validation and modification of a habitat suitability model for pronghorns. Wildlife Society Bulletin **13**:440-448.

Cunia, T. 1987. Construction of tree biomass tables by linear regression techniques. Pages 27-36 in Estimating tree biomass regressions and their error, USDA, Forest Service, General Technical Report NE-GTR-117.

Dunham, J. B., B. S. Cade, and J. W. Terrell. 2002. Influences of spatial and temporal variation on fish-habitat relationships defined by regression quantiles. Transactions of the American Fisheries Society **131**: 86-98.

Eastwood, P. D., G. J. Meaden, and A. Grioche. 2001. Modeling spatial variations in spawning habitat suitability for the sole *Solea solea* using regression quantiles and GIS procedures. Marine Ecology Progress Series **224**: 251-266.

Gerow, K., and C. Bilen. 1999. Confidence intervals for percentiles: an application to estimation of potential maximum biomass of trout in Wyoming streams. North American Journal of Fisheries Management **19**:149-151.

Gutenbrunner, C., J. Jurečková, R. Koenker, and S. Portnoy. 1993. Tests of linear hypotheses based on regression rank scores. Nonparametric Statistics **2**:307-331.

Haire, S. L., C. E. Bock, B. S. Cade, and B. C. Bennett. 2000. The role of landscape and habitat characteristics in limiting abundance of grassland nesting songbirds in an urban open space. Landscape and Urban Planning **48**(1-2):65-82.

Hubert, W. A., T. D. Marwitz, K. G. Gerow, N. A. Binns, and R. W. Wiley. 1996. Estimation of potential maximum biomass of trout in Wyoming streams to assist management decisions. North American Journal of Fisheries Management **16**:821-829.

Huston, M. A. 2002. Introductory essay: Critical issues for improving predictions. Pages 7-21 *in* J. M. Scott et al., editors. Predicting species occurrences: issues of accuracy and scale. Island Press, Covelo, California, USA.

Irwin, L. L., and J. G. Cook. 1985. Determining appropriate variables for a habitat suitability model for pronghorns. Wildlife Society Bulletin **13**:434-440.

Kaiser, M. S., P. L. Speckman, and J. R. Jones. 1994. Statistical models for limiting nutrient relations in inland waters. Journal of the American Statistical Association **89**:410-423.

Knight, C. A., and D. D. Ackerly. 2002. Variation in nuclear DNA content across environmental gradients: a quantile regression analysis. Ecology Letters **5**:66-76.

Koenker, R. 1994. Confidence intervals for regression quantiles. *Pages* 349-359 *in* P. Mandl and M. Hušková, editors. Asymptotic statistics: Proceedings of the 5th Prague Symposium. Physica-Verlag

Koenker, R., and G. Bassett. 1978. Regression quantiles. Econometrica **46**:33-50.

Koenker, R. and V. d'Orey. 1987. Computing regression quantiles. Applied Statistics

36:383-393.

Koenker, R., and O. Geling. 2001. Reappraising medfly longevity: a quantile regression survival analysis. Journal of the American Statistical Association **96**:458-468.

Koenker, R., and K. F. Hallock. 2002. Quantile regression: An introduction. Journal Economic Perspectives, In press.

Koenker, R., and J. A. F. Machado. 1999. Goodness of fit and related inference processes for quantile regression. Journal of the American Statistical Association **94**:1296-1310.

Koenker, R., and B. J. Park. 1996. An interior point algorithm for nonlinear quantile regression . Journal of Econometrics **71**:265-283.

Koenker, R., and F. Schorfheide. 1994. Quantile spline models for global temperature change. Climatic Change **28**:395-404.

Koenker, R., P. Ng, and S. Portnoy. 1994. Quantile smoothing splines. Biometrika **81**:673-680.

McClain, C. R., and M. A. Rex. 2001. The relationship between dissolved oxygen concentration and maximum size in deep-sea turrid gastropods: an application of quantile regression. Marine Biology **139**:681-685.

McCullagh, P., and J. A. Nelder. 1989. Generalized linear models. Chapman and Hall, New York.

Mosteller, F., and J. W. Tukey. 1977. Data analysis and regression. Addison-Wesley, New York.

Neter, J., M. H. Kutner, C. J. Nachtsheim, and W. Wasserman. 1996. Applied linear statistical models. Irwin, Chicago, Illinois, USA. 1408pp.

Portnoy, S. 1991. Asymptotic behavior of the number of regression quantile breakpoints. SIAM Journal of Science and Statistical Computing **12**:867-883.

Rosenbaum, P. R. 1995. Quantiles in nonrandom samples and observational studies. Journal of the American Statistical Association **90**:1424-1431.

Rosenbaum, P. R. 1999. Reduced sensitivity to hidden bias at upper quantiles in observational studies with dilated treatment effects. Biometrics 55:560-564.

Scharf, F. S., F. Juanes, and M. Sutherland. 1998. Inferring ecological relationships from the edges of scatter diagrams: comparison of regression techniques. Ecology **79**:448-460.

Terrell, J. W., B. S. Cade, J. Carpenter, and J. M. Thompson. 1996. Modeling stream fish habitat limitations from wedged-shaped patterns of variation in standing stock. Transactions of the American Fisheries Society **125**:104-117.

Thomson, J. D., G. Weiblen, B. A. Thomson, S. Alfaro, and P. Legendre. 1996. Untangling multiple factors in spatial distributions: Lilies, gophers, and rocks. Ecology **77**:1698-1715.

Vardeman, S. B. 1992. What about the other intervals? The American Statistician **46**: 193-197.

Welsh, A. H., R. J. Carroll, and D. Rupert. 1994. Fitting heteroscedastic regression models. Journal of American Statistical Association **89**:100-116.

Yu, K., and M. C. Jones. 1998. Local linear quantile regression. Journal of the American Statistical Association **93**:228-237.

Zhou, K. Q., and S. L. Portnoy. 1996. Direct use of regression quantiles to construct confidence sets in linear models. Annals of Statistics **24**:287-306.

Zhou, K. Q., and S. L. Portnoy. 1998. Statistical inference on heteroscedastic models based on regression quantiles. Nonparametric Statistics **9**:239-260.

Chapter 2

Rankscore and Permutation Testing Alternatives for Regression Quantile Estimates

Abstract: Performance of quantile rankscore tests used for hypothesis testing and constructing confidence intervals for linear quantile regression estimates ($0 \leq \tau \leq 1$) were evaluated for conditions relevant to ecological investigations of animal responses to their physical environment. Conditions evaluated included models with 2 - 6 predictors, moderate collinearity among predictors, homogeneous and heterogeneous errors, small to moderate samples (n = 20 - 300), and central to upper quantiles (0.50 - 0.99). Test statistics evaluated were the conventional quantile rankscore T statistic that is distributed as a Chi-square random variable with q degrees of freedom (where q parameters are constrained by H_0:) and an F statistic with its sampling distribution approximated by permutation or by an F distribution. The permutation F test maintained better Type I errors than the T test for models with smaller n, greater number of parameters p, and more extreme quantiles τ. Both versions of the test required weighting to maintain correct Type I errors when heterogeneity under the alternative model increased to around 5 standard deviations across the domain of X. A double permutation scheme was found to improve Type I errors for the permutation F test when null models were forced through the origin, as when testing the intercept or

any parameter in weighted models. Power was similar for conditions where both *T* and *F* tests maintained correct Type I errors. Confidence intervals on parameters and tolerance intervals for future predictions were constructed based on test inversion for an example application relating trout densities to stream channel width:depth.

1. Introduction

Estimating the quantiles ($0 \leq \tau \leq 1$) of a response variable conditional on some set of covariates in a linear model has many applications in the biological and ecological sciences (Terrell et al. 1996, Scharf et al. 1998, Cade et al. 1999, Cade and Guo 2000, Haire et al. 2000, Eastwood et al. 2001, Dunham et al. 2002). Quantile regression models allow the entire conditional distribution of a response variable y to be related to some covariates X, providing a richer description of functional changes than is possible by focusing on just the mean (or other central statistics), yet requiring minimal distributional assumptions (Koenker and Bassett 1978, 1982, Koenker and Machado 1999). Quantile regression estimates are especially enlightening for relationships involving heterogeneous responses where by definition rates of change are not the same across all parts of the response distribution.

Regression quantile models have been used where scientific considerations suggested that upper quantiles near the maximum better estimated effects of the biological process being measured as a limiting constraint (Cade et al. 1999, Cade and Guo 2000, Huston 2002). Statistical difficulties associated with characterizing limiting factors in ecology occur because the measured factor(s) may limit an organism only at some times or places, whereas other factors that were not measured may be limiting otherwise (Kaiser et al. 1994, Thomson et al. 1996, Cade et al. 1999, Cade and Guo 2000, Huston 2002). Temporal and spatial shifts in ecological limiting factors are to be expected. In an observational study it is impossible to know whether the measured covariates describe the factor actually limiting the organism at the time and location of sampling. Consequently, there may be large, unexplained heterogeneity in responses across levels of the measured covariates such that rates of change are less for some conditional central statistics (e.g., means or medians) compared to those for more extreme parts of the distribution (e.g., 90 - 99th percentiles). Heterogeneity induced by interaction effects of unmeasured but important processes (Cade et al. 1999, Huston 2002) creates a form of hidden bias typical in observational studies (Rosenbaum 1991, 1995).

Regression quantiles offer an estimation approach with considerable appeal both for prediction and understanding, regardless of whether interest is in extreme quantiles (e.g., 95 - 99th percentiles) for characterizing the boundary of a response distribution associated with some limiting factor (Cade and Guo 2000), or simply as a flexible method of estimating effects associated with heterogeneous distributions (Allen et al. 2001). Interpretations and properties of the estimated effects in regression quantiles are similar to more familiar linear modeling procedures such as least squares regression, but now are made for a family of quantiles in some interval that is selected based on scientific considerations (Cade et al. 1999, Koenker and Machado 1999, Koenker and Hallock In press). Regression quantile estimates also have a useful property not shared by estimates of means, equivariance under any monotonic transformation, that actually allows for simpler implementations and interpretations for transformable nonlinear models (Buchinsky 1995, Cade et al. 1999, Koenker and Geling 2001).

There is a well developed theory for estimating covariance matrices to provide inferences with asymptotic validity for linear regression quantile models (Koenker and Bassett 1978, 1982, Koenker and Machado 1999). These covariance methods rely on estimating the reciprocal of the error density function at the quantile of interest, $f(F^{-1}(0))$, i.e., the sparsity function. Performance of these asymptotic covariance methods at smaller sample sizes often is poor (Koenker 1987, Buchinsky 1991) and the asymptotic theory becomes suspect at more extreme (>0.7 and <0.3) quantiles (Chernozhukov and Umantsev 2001). Koenker (1994) introduced the idea of constructing confidence intervals by inverting a quantile rankscore test (Gutenbrunner et al. 1993) which does not require estimating the sparsity function and was expected to perform well under linear heteroscedastic regression models. The quantile rankscore test performed well at smaller sample sizes typically encountered in biological and ecological investigations in the limited simulations of Koenker (1994).

Questions remain about performance of the quantile rankscore test and potential modifications. In typical unimodal error distributions where density of the errors decreases as one moves away from the median, sampling variation and power of the more extreme quantiles (e.g., the 0.95 quantile) will be reduced compared to more central quantiles such as the median (0.50). How rapidly performance erodes will be a function of the error distribution, sample size, and number of parameters in a model. It was, thus, of interest to investigate performance of the quantile rankscore test across a range of quantiles, sample sizes, error distributions, and model structures to determine where inferences become unreliable. As the quantile of interest approaches 0 (minimum) or 1 (maximum), inferences may be more amenable to extreme value testing theory than conventional testing approaches (Chernozhukov and Umantsev 2001). The intercept parameter in a quantile regression model can be tested with the quantile rankscore test, although this clearly is excluded by the general theory of rankscore tests (Gutenbrunner et al. 1993). If the quantile rankscore test for the intercept parameter provides valid inferences, this procedure could be used for constructing confidence intervals at any specified value of the covariates. Extensions to prediction and tolerance intervals for some regression model forms would then be possible. Although the quantile rankscore test was evaluated for some linear heteroscedastic model forms and found to perform well (Koenker 1994), Koenker and Machado (1999) recently proposed a weighted modification of the quantile rankscore tests, where weights were a function of heterogeneity under the null hypothesis. A more systematic evaluation of the effects of heterogeneity on performance of the quantile rankscore test would help determine when it is desirable to use a weighted version of the test statistic.

Here I evaluated performance of the unweighted form of the quantile rankscore test for central to extreme quantiles, a range of error structures, small to moderate sample sizes, and model forms likely to be encountered in ecological applications where the objective is to estimate some organism’s response to its environment. Based on relationships between the asymptotic Chi-square form of the quantile rankscore test statistic and an F-test in a linear model, I considered alternative versions of rankscore tests that were evaluated by permutation arguments as well as by standard distributional theory. Weighted forms of the rankscore tests based on weighted quantile regression estimates also were evaluated. The alternative inference procedures were applied to a quantile regression analysis of Lahontan cutthroat trout (*Oncorhynchus clarki henshawi*) response to variations in their stream habitat, expanding on the previous analyses of Dunham et al. (2002).

2. Quantile Regression Model

The τ^{th} regression quantile ($0 \leq \tau \leq 1$) for the heteroscedastic linear location-scale model $\mathbf{y} = \mathbf{X\beta} + \mathbf{\Gamma\varepsilon}$ is defined as $Q_y(\tau|\mathbf{X}) = \mathbf{X\beta}(\tau)$ and $\boldsymbol{\beta}(\tau) = \boldsymbol{\beta} + F_\varepsilon^{-1}(\tau)\boldsymbol{\gamma}$; where $\mathbf{y}$ is an $n \times 1$ vector of dependent responses, $\boldsymbol{\beta}$ is a $p \times 1$ vector of unknown regression parameters, $\mathbf{X}$ is an $n \times p$ matrix of predictors (first column consists of 1's for an intercept term), $\boldsymbol{\gamma}$ is a $p \times 1$ vector of unknown scale parameters, $\mathbf{\Gamma}$ is a diagonal $n \times n$ matrix where the n diagonal elements are the n corresponding ordered elements of the $n \times 1$ vector $\mathbf{X\gamma}$ (diag($\mathbf{X\gamma}$)), $\boldsymbol{\varepsilon}$ is an $n \times 1$ vector of random errors that are independent and identically distributed (iid), and F_ε^{-1} is the inverse of the cumulative distribution of the errors (Koenker and Bassett 1982, Buchinsky 1991, Gutenbrunner and Jurečková 1992, Koenker and Machado 1999). Homoscedastic regression models are a special case of the linear location-scale model when $\boldsymbol{\gamma} = (1,0,...,0)'$ and $Q_y(\tau|\mathbf{X}) = \mathbf{X\beta}(\tau)$, $\boldsymbol{\beta}(\tau) = \boldsymbol{\beta} + (F_\varepsilon^{-1}(\tau),0,...,0)'$, where all parameters other than the intercept (β_0) in $\boldsymbol{\beta}(\tau)$ are the same for all τ. More general forms of heteroscedastic errors can be accommodated with regression quantiles (Koenker 1997, Koenker and Machado 1999) but were not considered here.

The restriction imposed on F_ε to estimate regression quantiles is that a τ^{th} quantile of **y - Xβ**(τ) conditional on **X** equals 0, $F_\varepsilon^{-1}(\tau|\mathbf{X}) = 0$. Estimates, **b**(τ), of **β**(τ) are solutions to the following minimization problem:

$$
\begin{gathered}
\min\Big[\sum_{i=1}^{n} \rho_\tau\big(y_i - \sum_{j=0}^{p} b_j x_{ij}\big)\Big] \\
\textit{where}\ \rho_\tau(e) = e(\tau - I(e<0)), \\
\textit{and}\ I(\cdot)\ \textit{is the indicator function}.
\end{gathered}
\qquad (1)
$$

The estimating equations in (1) yield primal solutions in a modification of the Barrodale and Roberts (1974) simplex linear program for any specified value of τ (Koenker and d'Orey 1987). With little additional computation the entire regression quantile process for all distinct values of τ can be estimated (Koenker and d'Orey 1987, 1994).

Consistent estimates with reduced sampling variation for heteroscedastic linear models can be obtained by implementing weighted versions of the regression quantile estimators, where weights are based on the sparsity function at a given quantile and covariate value (Koenker and Portnoy 1996, Koenker and Machado 1999). In the linear location-scale model this simplified to using an $n \times n$ weights matrix, $\mathbf{W} = \mathbf{\Gamma}^{-1}$, where the $p \times 1$ vector of scale parameters $\boldsymbol{\gamma}$ would usually have to be estimated in applications (Gutenbrunner and Jurečková 1992, Koenker and Zhao 1994, Koenker and Machado 1999). The weighted regression quantile estimates then are given by

$$\begin{gathered}\min\left[\sum_{i=1}^{n}\rho_\tau\left(y_i-\sum_{j=0}^{p}b_j x_{ij}\right)w_i\right]\\ \textit{where}\ \ \rho_\tau(e)=e(\tau-I(e<0)),\\ w_i\ \textit{is a weight,}\\ \textit{and}\ I(\cdot)\ \textit{is the indicator function.}\end{gathered} \tag{2}$$

which is easily implemented by multiplying **y** and **X** by **W** and then using the unweighted estimator (1).

3. Rankscore Test Statistics

The primal linear programming solution for (1) has as its corresponding dual solution

$$\begin{gathered}\max\{y'a \mid X'a = (1-\tau)X'1,\ a\in[0,1]^n\}\\ \textit{where}\ 1\ \textit{denotes an}\ n\textit{-vector of}\ 1\textit{'s}\end{gathered} \tag{3}$$

that serves as the basis for constructing rankscore tests using the regression quantile estimates (Gutenbrunner et al. 1993, Koenker and d'Orey 1994, Koenker 1994, 1997). The τ-quantile rankscore test uses the τ-quantile score function, $\varphi_\tau(t) = \tau - I(t < \tau)$, on the $n \times 1$ vector of dual linear programming solutions, $\mathbf{a}(\tau) = [0,1]^n$, associated with estimating the reduced parameter model corresponding to constraints imposed by the null hypothesis on the full parameter model. The reduced parameter model, $\mathbf{y} - \mathbf{X}_2\xi(\tau) = \mathbf{X}_1\boldsymbol{\beta}_1(\tau) + \boldsymbol{\Gamma\varepsilon}$, is constructed by partitioning $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2)$, where $\mathbf{X}_1$ is $n \times (p - q)$ and $\mathbf{X}_2$ is $n \times q$; and by partitioning $\boldsymbol{\beta} = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2)$, where $\boldsymbol{\beta}_1(\tau)$ is a $(p - q) \times 1$ vector of unknown nuisance parameters under the null and $\boldsymbol{\beta}_2(\tau)$ is a $q \times 1$ vector of parameters specified by the null hypothesis H_0: $\boldsymbol{\beta}_2(\tau) = \xi(\tau)$ (frequently $\xi(\tau) = \mathbf{0}$) for the full parameter model $\mathbf{y} = \mathbf{X}_1\boldsymbol{\beta}_1(\tau) + \mathbf{X}_2\boldsymbol{\beta}_2(\tau) + \boldsymbol{\Gamma\varepsilon}$; and $\mathbf{y}$, $\boldsymbol{\Gamma}$, and $\boldsymbol{\varepsilon}$ are as above. The $n \times 1$ vector of rankscores $\mathbf{r}(\tau) = \mathbf{a}(\tau) - (1 - \tau)\mathbf{1}$, where $\mathbf{1}$ denotes an $n \times 1$ vector of 1's, is regressed on the design matrix and the test statistic

$$T = \mathbf{S}(\tau)'\mathbf{Q}^{-1}\mathbf{S}(\tau)/(\tau(1 - \tau)), \tag{4}$$

where $\mathbf{Q} = n^{-1}\mathbf{X}_2'(\mathbf{I} - \mathbf{X}_1(\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1')\mathbf{X}_2$ and $\mathbf{S}(\tau) = n^{-0.5}(\mathbf{X}_2 - \mathbf{X}_1(\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{X}_2)'\mathbf{r}(\tau)$, is asymptotically distributed under H_0: as χ^2 with q degrees of freedom. The elements of $\mathbf{a}(\tau)$ are 1 when the residuals for the reduced model are positive, 0 when the residuals are negative, and in the interval (0,1) when the residuals are 0, i.e., observations fit exactly by the τ^{th} regression quantile estimate. Rankscores $\mathbf{r}(\tau)$ then are τ for positive residuals, τ - 1 for negative residuals, and in the interval (τ - 1, τ) when residuals are 0. The rankscores, $\mathbf{r}(\tau)$, correspond to the quantile weights used in estimating the reduced parameter null model in (1). Validity of the rankscore test assumes a positive density for y at the estimate, $f(F^{-1}(\tau)) > 0$.

If $\mathbf{X}_2 = \mathbf{x}_2$ and $\boldsymbol{\beta}_2(\tau)$ is a scalar, i.e., a single predictor is being tested, then the quantile rankscore statistic simplifies and under the null hypothesis this 1 degree of freedom test is referenced to a standard normal distribution (Koenker 1994, 1997). This construction allows confidence intervals to be easily estimated by inversion with a modification of the linear program used to estimate regression quantiles (Koenker 1994). Because the sampling distribution of the rankscore test statistic is discontinuous, Koenker (1994) recommended interpolating between adjacent hypothesized values of $\boldsymbol{\beta}_2(\tau) = \xi(\tau)$ for constructing confidence intervals when inverting quantile rankscore tests. Confidence intervals estimated by inverting the quantile rankscore test may be asymmetric.

The τ-quantile rankscore test is based on a nondecreasing, square integrable scoring function with mean $\mu(\varphi) = 0$ and variance $\sigma^2(\varphi) = \tau(1- \tau)$ and, thus, is similar in form to the aligned rank transform statistic considered by Mansouri (1999). Note that $\mathbf{S}(\tau)'\mathbf{Q}^{-1}\mathbf{S}(\tau)$ in (4) is the sum of squares of regression for $\mathbf{X_2}$, SSReg$(\tau) =$ SSE$(\tau)_{\text{red}}$ - SSE$(\tau)_{\text{full}}$, where SSE$(\tau)_{\text{red}} = \mathbf{r}(\tau)'(\mathbf{I} - \mathbf{X_1}(\mathbf{X_1}'\mathbf{X_1})^{-1}\mathbf{X_1}')\mathbf{r}(\tau)$ and SSE$(\tau)_{\text{full}} =$ $\mathbf{r}(\tau)'(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{r}(\tau)$. Mansouri (1999) proved that a test statistic form like (4) was just the limiting ($n \rightarrow \infty$) form of an F statistic,

$$F_{q,\, n - p} = (\text{SSE}(\tau)_{\text{red}} - \text{SSE}(\tau)_{\text{full}})/(q\text{MSE}(\tau)), \tag{5}$$

where MSE(τ) = SSE$(\tau)_{\text{full}}/(n - p) \rightarrow \sigma^2(\varphi)$, and established via simulation that (5) had better small sample Type I error rates than (4). Because the sampling distribution of the τ-quantile rankscore test is discontinous and increases in discontinuity as τ approaches 0 or 1, I expected that there might be some small sample performance advantages to using (5) over (4) for hypothesis tests or constructing confidence intervals by inverting the quantile rankscore test.

The F statistic for the quantile rankscore test (5) is based on a regression with a dependent variable, $\mathbf{r}(\tau)$, that is a function of residuals under the reduced parameter null model. This test statistic is amenable to evaluation by permutation arguments that have been developed for testing subhypotheses in least squares regression (Kennedy and Cade 1996, Anderson and Legendre 1999, Anderson and Robinson 2001). The permutation distribution computed for (5) might yield more reliable Type I error rates at smaller sample sizes and more extreme quantiles than the F distribution approximation with q and $n - p$ degrees of freedom. The quantile rankscore F test evaluated by permutation arguments is defined by the slightly simpler form for the observed value of the statistic

$$F_o = (\text{SSE}(\tau)_{\text{red}} - \text{SSE}(\tau)_{\text{full}})/(\text{SSE}(\tau)_{\text{full}}), \tag{6}$$

where $\text{SSE}(\tau)_{\text{red}}$ and $\text{SSE}(\tau)_{\text{full}}$ are as above, because the degrees of freedom in (5) are unnecessary as they are invariant under permutation. Note that $F_o \times \text{SSE}(\tau)_{\text{full}}/(\tau(1 - \tau)) = T$ and $F_o \times ((n - p)/q) = F_{q,\, n - p}$. The permutation test statistic, F_o, has a simple interpretation as a proportionate reduction in sums of squares when passing from reduced to full parameter models for a specified quantile.

Following Kennedy and Cade (1996), Anderson and Legendre (1999), and Anderson and Robinson (2001), the observed value of the rankscore test statistic, F_o, is evaluated under the null hypothesis by permuting the τ-quantile rankscores, $\mathbf{r}(\tau)$, among the rows of the design matrix ($\mathbf{X}$) with equal probability, $(n!)^{-1}$. A large random sample of size m is used to approximate the $n!$ possible permutations. Probability under the null hypothesis that $F \geq F_o$ is approximated by (the number of $F \geq F_o$ + 1)/(m + 1). I used a minimum of $m + 1 = 10{,}000$ to achieve probability approximations with minimal variation due to the Monte Carlo resampling.

Although permuting residuals ($\mathbf{e} = \mathbf{y} - \mathbf{X}_1\mathbf{b}_1$) under the null reduced parameter model does not in general yield exact permutation probabilities except when the null parameter is just an intercept (β_0), this permutation approach due to Freedman and Lane (1983) was found to have perfect correlation asymptotically with the exact test (only possible when $\boldsymbol{\beta}_1$ is known) (Anderson and Robinson 2001) and has performed well in simulation studies (Cade and Richards 1996, Kennedy and Cade 1996, Anderson and

Legendre 1999, Legendre 2000). There is some correlation ($-(n - 1)^{-1}$) among the residuals and they don't have constant variance ($E[\mathbf{e}\,\mathbf{e}'] = \sigma^2(\mathbf{I} - \mathbf{X_1}(\mathbf{X_1}'\mathbf{X_1})^{-1}\mathbf{X_1}')$), implying that they are not exactly exchangeable. Dependency among the residuals decreases with increasing sample size providing some asymptotic justification for treating them as exchangeable random variables (Randles 1984). Commenges (In Press) established that transformations to preserve exchangeability of the first two moments of the residuals must reduce the rank of the $n \times 1$ vector of residuals to an $(n - p + q) \times 1$ vector of uncorrelated residuals, e.g., the best linear unbiased residuals with scalar covariance of Theil (1965). This approach was not pursued here. However, the τ-quantile rankscore transformation of residuals to $[\tau - 1, \tau]$ under the null model should approach constant variance more rapidly than raw residuals. There are at most $n - p + q$ residuals with rankscores of τ or $\tau - 1$, and at least $p - q$ rankscores in the interval $(\tau - 1, \tau)$. Together these conditions should reduce dependency among the transformed residuals and improve exchangeability under the null model.

An obvious modification of the quantile rankscore tests *T* and *F* is to incorporate a weights matrix, **W**, in estimating the reduced parameter null model and in constructing the test statistics (4), (5) and (6). The disadvantage of this approach is that in applications the weights are unknown and must be estimated. Part of the motivation for the quantile rankscore test was a belief that converting to scores in the interval $[\tau - 1, \tau]$ would eliminate the need to formally model error heterogeneity (Koenker 1994). Koenker and Machado (1999) proposed a weighted version of the quantile rankscore test, where weights were a function only of predictors in the null model ($\mathbf{X_1}$), whereas I considered weights as a function of predictors in the full model (**X**).

4. Simulation Experiment

Although my primary interest was in performance of the test statistics for regression quantile models estimated with heterogeneous responses, I first conducted a set of Monte Carlo simulations with homogeneous errors to establish performance for models with simpler error structure. Normal (μ = 0, σ = 1), uniform (min = -2, max = 2), and lognormal (median = 0, σ = 0.75) error distributions were used to provide responses with symmetric, unimodal variation with greatest density at the center, symmetric variation with constant density, and asymmetric variation with density in a long upper tail. A limited set of simulations with Poisson error distributions was made to evaluate the quantile rankscore test when there were many tied integer values as would occur with counts of organisms, violating the assumption of positive density at the estimates. Error distributions were centered on their 0.50, 0.75, 0.90, 0.95, or 0.99 quantiles so that $F_{\varepsilon}^{-1}(\tau|\mathbf{X}) = 0$, providing a range of central to extreme regression quantiles. Note that similar simulation results for quantiles in the lower tail (0.25, 0.10, 0.05, and 0.01) would be obtained for the symmetric error distributions.

Simple 2 parameter and 6 parameter multiple regression models were simulated for n = 20, 30, 60, 90, 150, and 300. Independent variables were structured to have a range of values and correlation structure similar to what might be expected in measures of forest habitat structure for avian species. Independent variables were structured so that X_0 was a column of 1's for the intercept; X_1 was uniformly distributed (0, 100); X_2 was negatively correlated (r = -0.89) with X_1 specified by the function X_2 = 4,000

$-20X_1 + N(\mu = 0, \sigma = 300)$; X_3 was positively correlated ($r = 0.94$) with X_1 specified by the function $X_3 = 10 + 0.4X_1 + N(\mu = 0, \sigma = 16)$; X_4 was a 0,1 indicator variable randomly assigning half the sample to each of 2 groups; and X_5 was the multiplicative interaction of X_3 and X_4. Thus, X_1 ranged from 0 - 100 similar to measures of percent tree canopy cover, X_2 had most values in the range 0 - 5,000 and was inversely related to tree cover similar to density (stems/ha) of a shade intolerant shrub, and X_3 had most values in the range 0 - 60 similar to tree height (m) and was positively related to tree cover. Variables X_2 and X_3 were negatively correlated ($r = -0.85$) with each other through their indirect functional relation with X_1. The indicator variable (X_4) and its interaction with X_3 (X_5) allowed the effect of X_3 for the regression quantile function to differ in slopes, intercepts, or both terms for the 2 groups.

Each combination of conditions (quantile, error distribution, sample size, and model structure) was sampled 1,000 times, and the test statistics T and F_o were computed for each sample. Probabilities for the permutation F test were evaluated with separate $m + 1 = 10{,}000$ random samples of the permutation distribution. Cumulative distribution function (cdf) plots of the Type I error probabilities under the null hypothesis were graphed and compared with the expected uniform cdf. However, point estimates for $\alpha = 0.05$ and 0.10, corresponding to coverage for 95% and 90% confidence intervals, were graphed across the combination of model conditions because the number of graphs required to display the cdf plots was excessive. The 99% binomial confidence interval for 1,000 simulations for $\alpha = 0.10$ is 0.076 - 0.124 and for $\alpha = 0.05$ is 0.032 - 0.068, which can be used as a guide to judge how much the estimated error rates exceeded variation expected from the sampling simulations. Power under the alternative hypotheses was graphed only for $\alpha = 0.05$ across all combinations of conditions, although cdf plots were initially examined.

All data for the simulation studies were generated with functions in S-Plus 2000 (Mathsoft, Inc., Seattle, WA). Regression quantile estimates and test statistics were computed by a static memory compilation of Fortran 95 routines implemented in the Blossom software available from the U. S. Geological Survey (*www.mesc.usgs.gov/products/software/ blossom.shtml*). Regression quantile estimates and *T* rankscore tests from the software used in simulations were compared with estimates from the S-Plus scripts developed by R. Koenker (*www.econ.uiuc.edu~roger/research/home.html*) for selected models both before and after simulations were completed and found to agree to at least 7 decimal places.

4.1 Homogeneous Error Structure - Simple Regression

The simple 2 parameter regression model, $y = \beta_0 + \beta_1 X_1 + \varepsilon$ was evaluated for H_0: $\beta_1 = 0$ with β_0 fixed at 6.0 and β_1 = 0.0, 0.01, 0.05, 0.10, and 0.20. Estimated Type I error rates ($\beta_1 = 0.0$) for the permutation *F* test maintained nominal rates across all conditions whereas the *T* test became excessively conservative for the 0.95 quantile for $n \le 30$ and for the 0.99 quantile for $n \le 150$ (Fig. 2.1). Results for the permutation test were consistent with exact exchangeability for this hypothesis. Type I errors for the 0.75 quantile were nearly identical to those for the 0.50 quantile and, therefore, were not graphed for this or subsequent simulations. Results were similar for all error distributions for most conditions so only results of the lognormal error distribution are

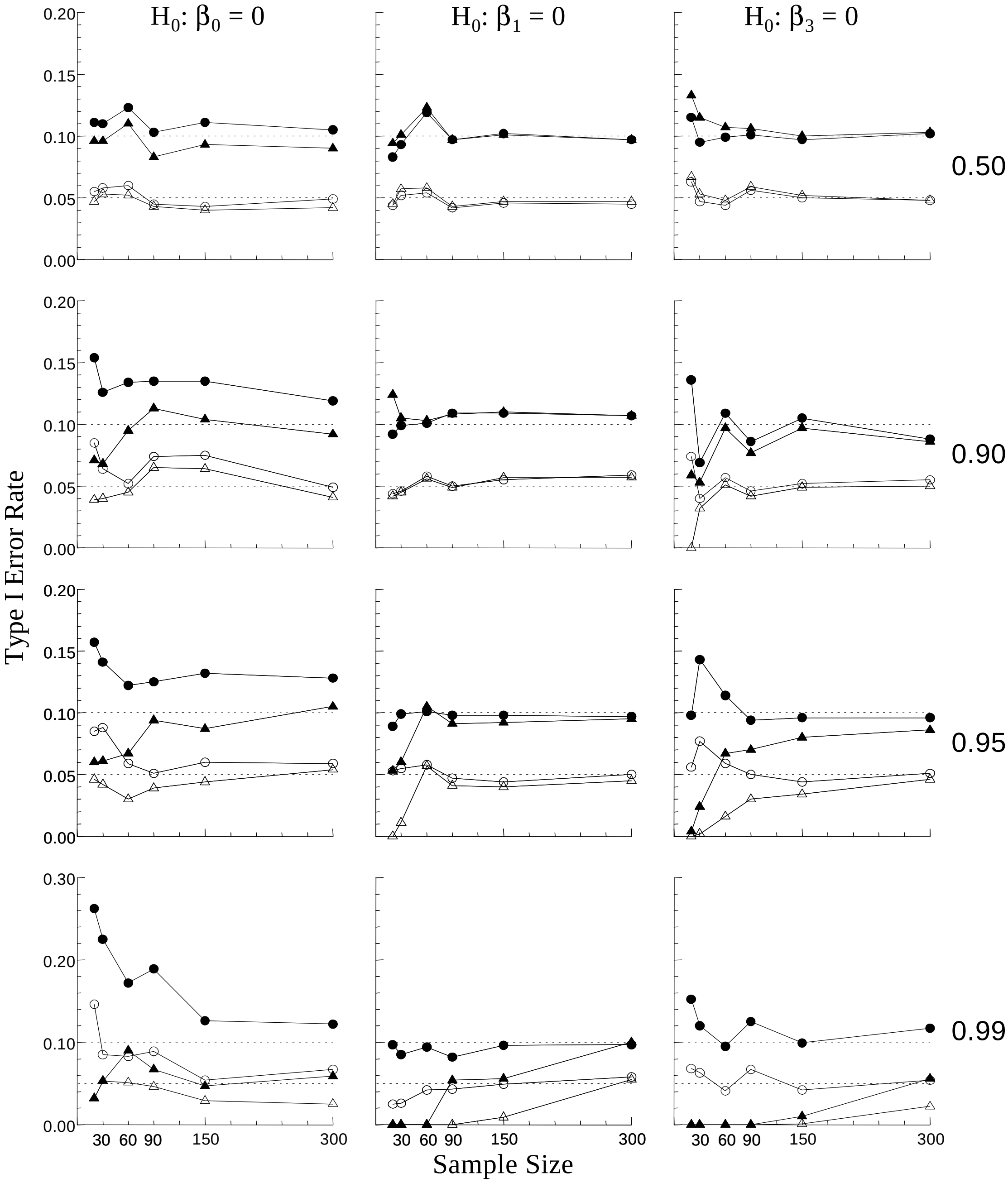


Figure 2.1. Estimated type I error rates for $\alpha = 0.05$ (open) and 0.10 (solid); for the permutation F (circles) and Chi-square distributed T (triangles) rankscore tests; for homogeneous lognormal error distributions; for H_0: $\beta_0 = 0$ and H_0: $\beta_1 = 0$ in the model $y = \beta_0 + \beta_1 X_1 + \varepsilon$, and H_0: $\beta_3 = 0$ in the model $y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \varepsilon$; for 0.50, 0.90, 0.95, and 0.99 quantiles; and for n = 20, 30, 60, 90, 150, and 300. 1,000 random samples were used at each combination of H_0:, n, and quantile.

given in the Figures for this and subsequent simulations. Results for normal and uniform error distributions are in Appendix 2.

It was possible to convert F_o to the $F_{q,\,n\,-\,p}$ rankscore statistic and evaluate probabilities with an F distribution with q and n - p degrees of freedom. The F distribution approximation controlled Type I errors under the same conditions where the Chi-square approximation of the T test statistic was well behaved and provided some improvement for smaller samples and more extreme quantiles. However, the F distribution did not maintain Type I errors as well as the permutation approximation at small n and more extreme quantiles. An example for the 0.99 quantile and lognormal error distribution demonstrates that the permutation F test had less discontinuous probabilities that were more uniformly distributed than those for the distributional approximations of T and $F_{q,\,n\,-\,p}$ (Fig. 2.2).

The simple 2 parameter regression model also was evaluated for H_0: $\beta_0 = 0$ with β_1 fixed at 0.10 and β_0 = 0.0, 0.5, 1.0, 2.0, and 3.0. Type I error rates for the intercept under then null hypothesis (β_0 = 0.0) were better maintained by the T test than the F test, which was always slightly liberal although not excessively so until 0.95 and 0.99 quantiles and n < 150 (Fig. 2.1). The T test was slightly conservative for the 0.95 quantile for n < 90 and for the 0.99 quantile for n < 300.

Detailed exploration of the simulation results for the permutation F test indicated that there was additional sampling variation not accounted for by the permutation distribution of the test statistic when the null model was constrained through the origin. If the number of positive, negative, and zero residuals are denoted

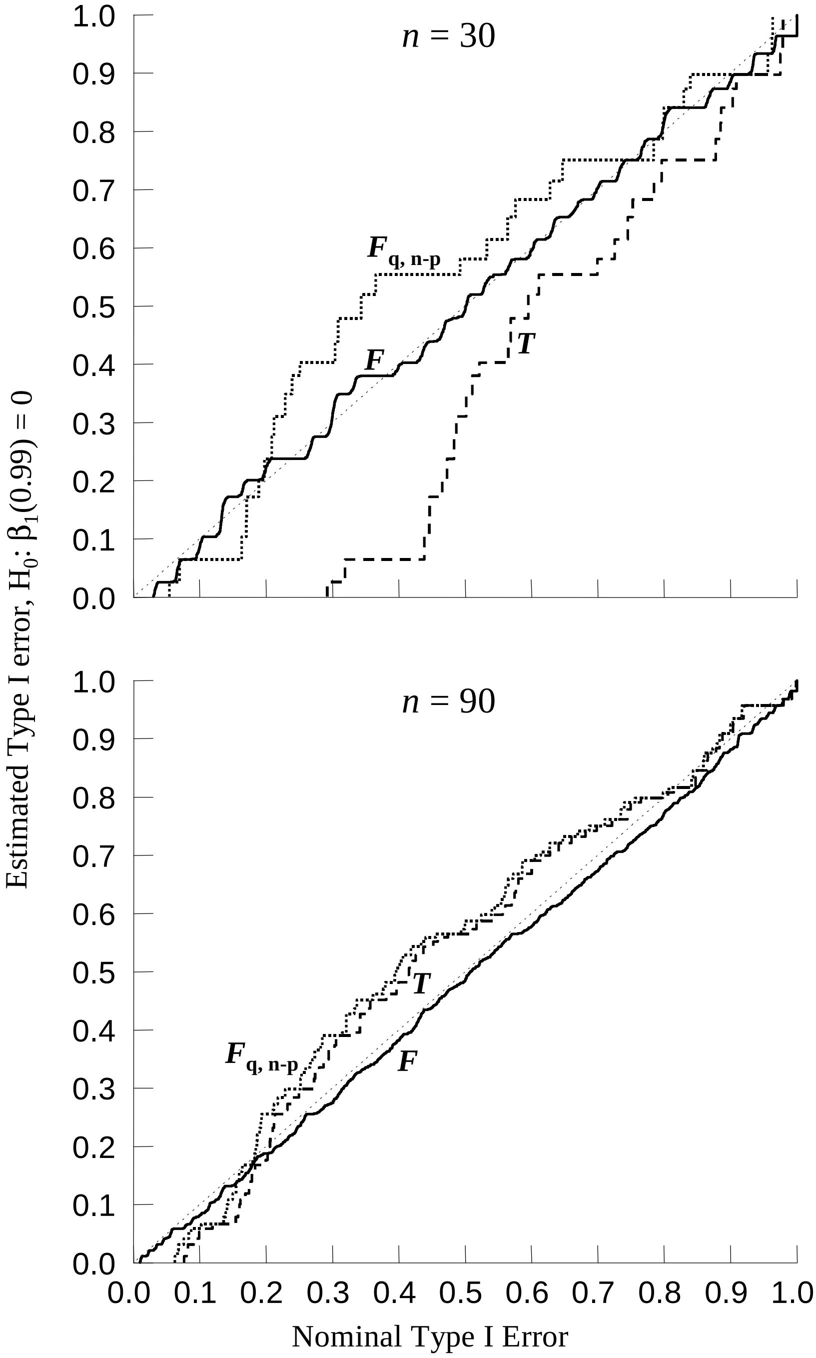


Figure 2.2. Cumulative distributions of 1,000 estimated errors for permutation approximation of the F (solid line) rankscore, $F_{q,\, n-p}$ distribution approximation of (square dot) the rankscore, and Chi-square distribution approximation of T rankscore tests (dashes) for H_0: $\beta_1 = 0$ for the 0.99 quantile, for n = 30 and 90, for the lognormal error distribution in the model $y = \beta_0 + \beta_1 X_1 + \varepsilon$.

by N^+, N^-, N^0, respectively, and if $N^0 = p - q$ under the null model, then there are at most $n\tau$ negative residuals ($N^- \leq n\tau \leq N^- + N^0$) and at most $n(1 - \tau)$ positive residuals ($N^+ \leq n[1 - \tau] \leq N^+ + N^0$) when the null model includes an intercept (Koenker and Bassett 1978, Koenker and Portnoy 1996). When the null model does not include an intercept, the limits on the number of positive (negative) residuals exceeded these values by amounts consistent with binomial random variation with success probability 1 - τ (or τ for negative residuals). Consequently, I modified a recently proposed double permutation scheme for least squares regression through the origin (Legendre and Desdevises In Press) for the quantile rankscore test as a possible remedy. The values of the rankscores, $\mathbf{r}(\tau)$, rather than being fixed across all permutations to $\mathbf{X}$ were varied such that the number of $\mathbf{r}(\tau)$ with value τ for positive residuals (and conversely values of $\tau - 1$ for negative residuals) was a binomial random variable with parameter $1 - \tau$. The double permutation F test for the intercept had improved Type I error rates that were similar to the T test when n was not too small and $\tau < 0.99$ but became excessively conservative when $\tau = 0.99$ and $n < 300$ (Fig. 2.3).

Power for nonzero slopes (β_1 = 0.01, 0.05, 0.10, 0.20) was similar for the F and T tests with a small improvement for the F test (relative power = 0.98 - 1.35) at 0.90 and 0.95 quantiles at smaller n (Fig.2.4). The F test provided effective power down to $n = 30$ for 0.95 and $n = 150$ for 0.99 quantiles, whereas the T test only provided effective power down to $n = 60$ and $n = 300$, respectively, because of very conservative Type I error rates at smaller sample sizes (Fig. 2.4). The drop in power when moving

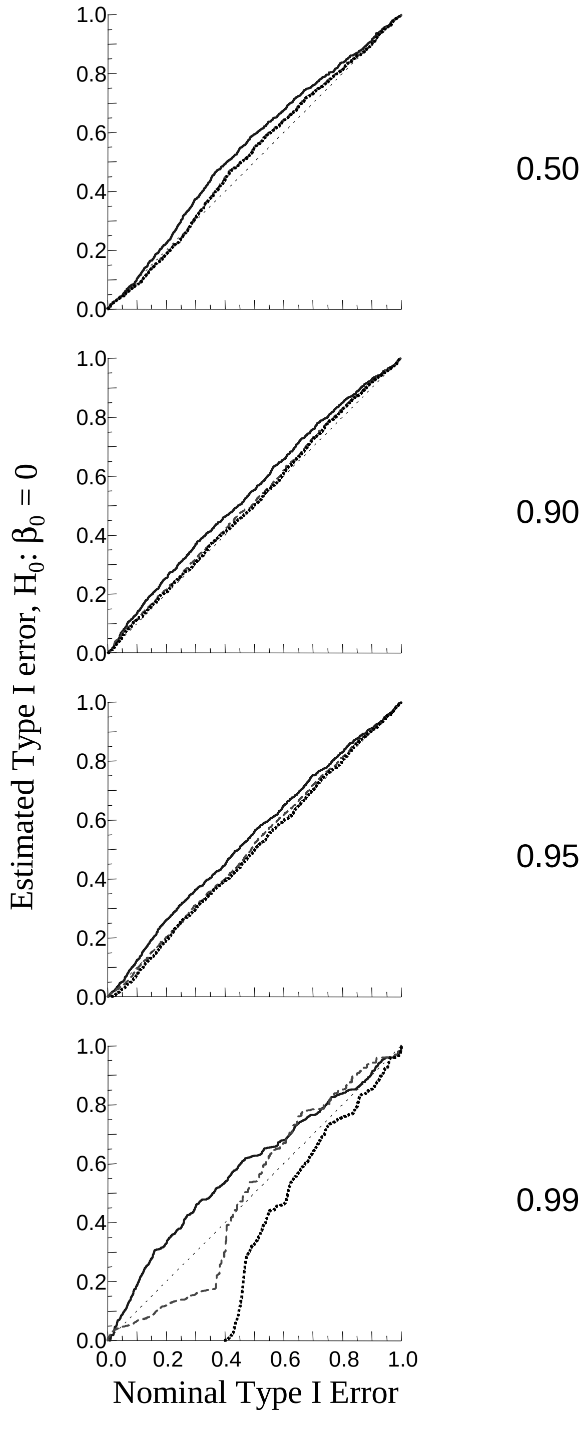


Figure 2.3. Cumulative distributions of 1,000 estimated errors for permutation *F* (solid), double permutation *F* (square dots), and *T* (dashed) rankscore tests of H_0: $\beta_0 = 0$; for 0.50, 0.90, 0.95, and 0.99 quantiles; for the model $y = \beta_0 + \beta_1 X_1 + \varepsilon$; for the lognormal error distribution and $n = 90$.

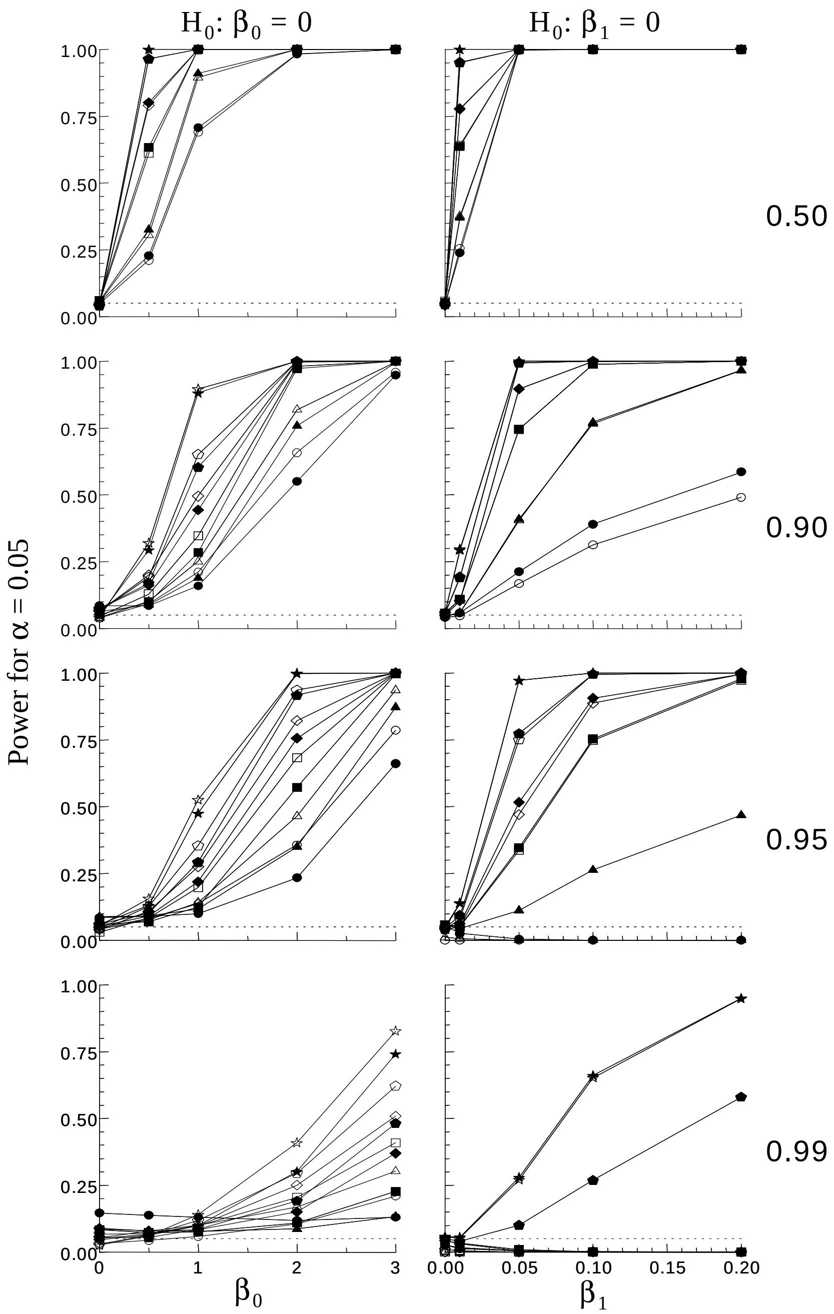


Figure 2.4. Estimated power for $\alpha = 0.05$ for the permutation F (solid) and Chi-square distributed T (open) rankscore tests; for homogeneous lognormal error distributions; for H_0: $\beta_0 = 0$ and H_0: $\beta_1 = 0$ in the model $y = \beta_0 + \beta_1 X_1 + \varepsilon$; for $\beta_0 = 0.0, 0.5, 1.0, 2.0$, and 3.0 and for $\beta_1 = 0.0, 0.01, 0.05, 0.10$, and 0.20; for 0.50, 0.90, 0.95, and 0.99 quantiles; and for $n = 20$ (circle), 30 (triangle), 60 (square), 90 (diamond), 150 (pentagon), and 300 (star). Open symbols often are hidden behind solid symbols when equal. 1,000 random samples were used at each combination of effect size, n, and quantile.

from 0.50 to 0.99 quantiles was greatest for the lognormal (Fig.2.4) and normal error distributions (Appendix 2.3) and least for the uniform error distribution (Appendix 2.4), which had a slight increase in power with increasing quantiles up to 0.95 for $n \geq 60$. Power for nonzero intercepts (β_0 = 0.5, 1.0, 2.0, and 3.0) was slightly greater for the *T* compared to the *F* test (Fig. 2.4). Similar results were obtained for the normal and uniform error distribution (Appendices 2.3 and 2.4). Power was not estimated for the double permutation *F* test but should be similar to the *T* test.

4.2 Homogeneous Error Structure - Multiple Regression

The 6 parameter model, $y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \varepsilon$, was evaluated for H_0: $\beta_3 = 0$ with $\beta_0 = 36.0$, $\beta_1 = 0.10$, $\beta_2 = -0.005$, $\beta_4 = 2.0$, and $\beta_3 = \beta_5 = 0.0$. The permutation *F* test maintained better Type I error rates for smaller *n* for the 0.95 and 0.99 quantiles than the *T* test (Fig. 2.1). The 6 parameter model also was evaluated for H_0: $\beta_4 = 0$ with $\beta_0 = 36.0$, $\beta_1 = 0.10$, $\beta_2 = -0.005$, $\beta_3 = 0.05$, and $\beta_4 = \beta_5 = 0.0$. Type I error rates were similar to those for H_0: $\beta_3 = 0$. Power was not investigated for multiple regression models with homogeneous errors.

4.3 Heterogeneous Error Structure - Simple Regression

The 2 parameter regression model with heterogeneous errors, $y = \beta_0 + \beta_1 X_1 + (1 + \gamma X_1)\varepsilon$, was evaluated with γ = 0.025, 0.05, and 0.10 for H_0: $\beta_1 = 0$ with $\beta_0 = 6.0$ and $\beta_1 = 0.0$ to evaluate the effects of increasing heterogeneity on Type I error rates for the rankscore tests. Type I error rates were increasingly liberal for the permutation *F* and *T* tests (Fig. 2.5) with increasing heterogeneity, except that the *T* test became excessively conservative at $n < 60$ for the 0.95 and at $n < 150$ for the 0.99 quantile. Results were

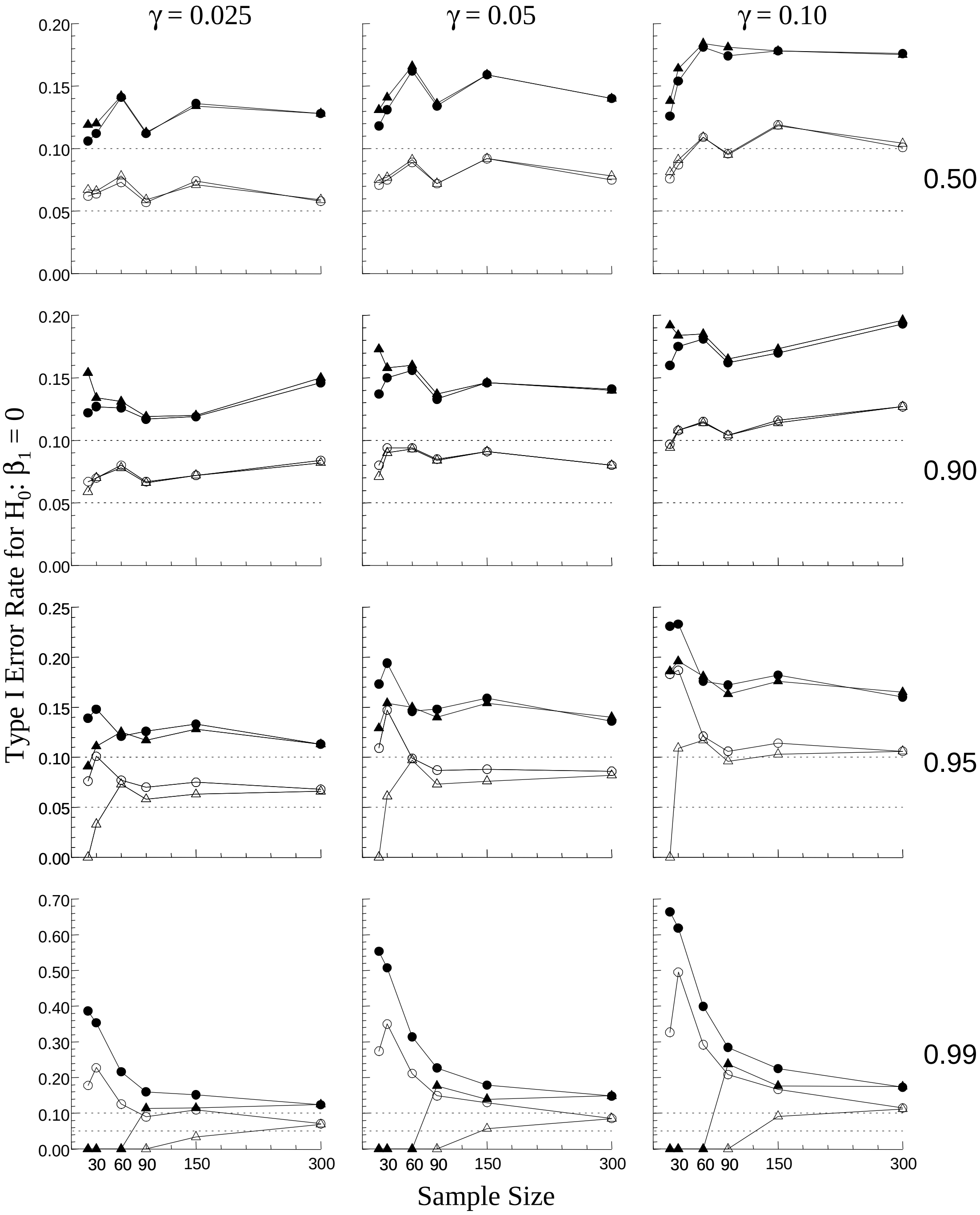


Figure 2.5. Estimated type I error rates for $\alpha = 0.05$ (open) and 0.10 (solid); for the permutation F (circles) and Chi-square distributed T (triangles) rankscore tests for H_0: $\beta_1 = 0$; for heterogeneous lognormal error distributions with $\gamma = 0.025$, 0.05, and 0.10 in the model $y = \beta_0 + \beta_1 X_1 + (1 + \gamma X_1)\varepsilon$; for 0.50, 0.90, 0.95, and 0.99 quantiles; and for n = 20, 30, 60, 90, 150, and 300. 1,000 random samples were used at each combination of γ, n, and quantile.

similar for normal and uniform error distributions (Appendices 2.5 and 2.6). Again, converting F_o to an $F_{q,\, n-p}$ rankscore statistic and evaluating probabilities with the F distribution and q and $n - p$ degrees of freedom provided minor improvements in small sample Type I errors compared to the T test, similar to simulations with homogeneous errors. The F distributional approximation did not maintain Type I error levels as well as the permutation approximation for the F test with small n and extreme quantiles. Type I error rates when $\gamma = 0.10$, which corresponds to a 10-fold increase in σ across the domain of X_1 since X_1 ranges 0 - 100, were such that nominal 95% confidence intervals would have actual coverage of only 90%.

Weighted versions of the regression quantile estimates and the rankscore tests for $\gamma = 0.05$ were simulated using the known weights, $w = (1 + 0.05X_1)^{-1}$, in (2). Type I error rates were improved for the weighted versions of both tests (Fig. 2.6B) compared to those for the unweighted tests (Fig. 2.5), except for the 0.99 quantile and smaller n. The permutation F test was always slightly more liberal than the T test because the weighted estimate for the null model is forced through the origin. Here, again the double permutation F test provided improved Type I errors over the permutation F test except at the 0.99 quantile and $n < 300$ (Fig. 2.6A), where none of the weighted statistics worked well.

The H_0: $\beta_0 = 0$ also was evaluated in the 2 parameter regression model with heterogeneous errors, $y = \beta_0 + \beta_1 X_1 + (1 + \gamma X_1)\varepsilon$, with $\gamma = 0.05$, $\beta_1 = 0.10$, and $\beta_0 = 0.0$, 0.5, 1.0, 2.0, and 3.0. The T test maintained Type I error rates ($\beta_0 = 0.0$) slightly better than the permutation F test similar to simulations for homogenous errors, with error

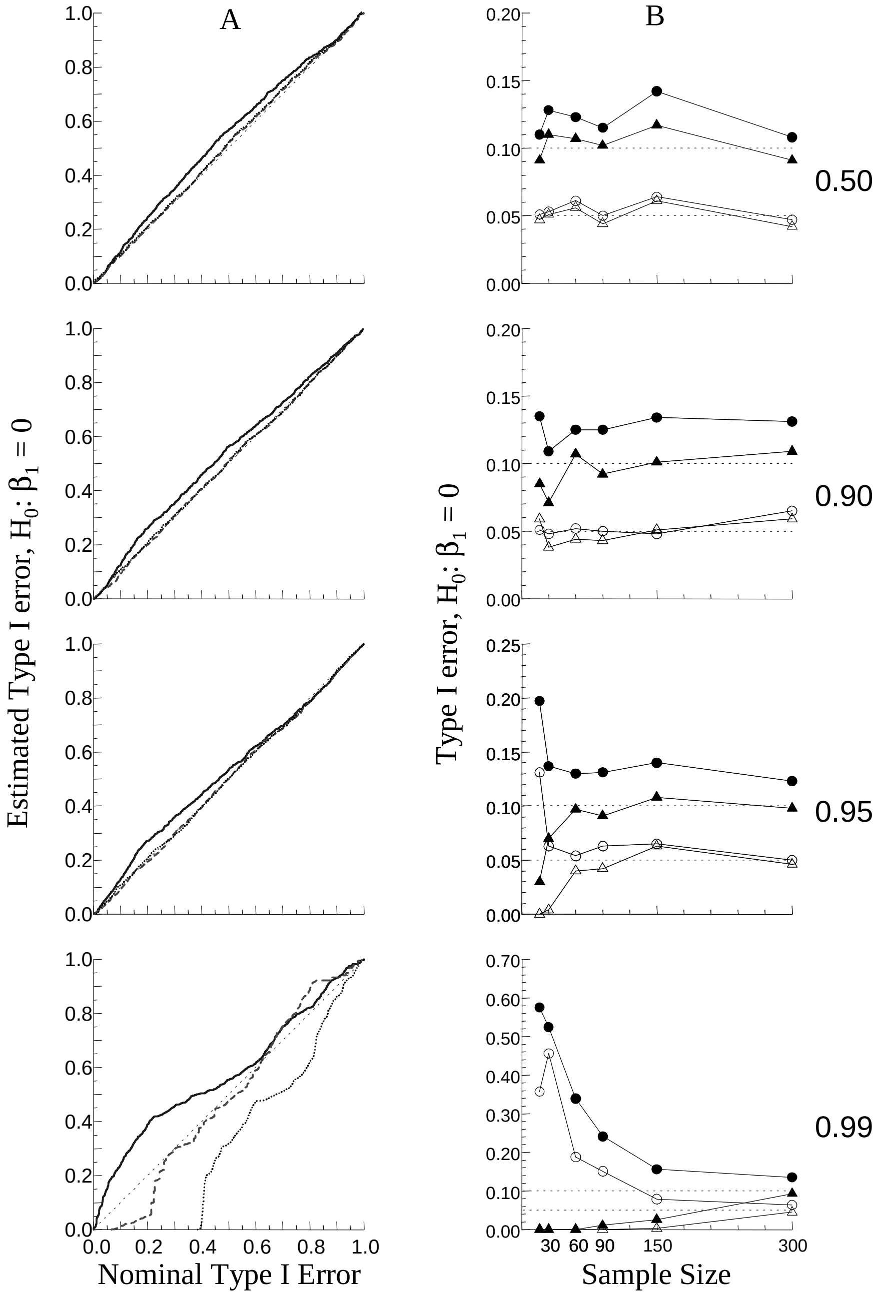


Figure 2.6. (A) Cumulative distributions of 1,000 estimated errors for permutation *F* (solid), double permutation *F* (square dots), and *T* (dashed) rankscore tests of H_0: $\beta_1 = 0$; for 0.50, 0.90, 0.95, and 0.99 quantiles; for the weighted model $wy = w\beta_0 + w\beta_1 X_1 + w(1 + \gamma X_1)\varepsilon$ with $\gamma = 0.05$ and $w = (1 + \gamma X_1)^{-1}$; for the lognormal error distribution and $n = 90$. (B) Estimated type I error rates for $\alpha = 0.05$ (open) and 0.10 (solid); for the permutation *F* (circles) and Chi-square distributed *T* (triangles) rankscore tests for H_0: $\beta_1 = 0$; for the same weighted model with lognormal error distributions; for 0.50, 0.90, 0.95, and 0.99 quantiles; and for n = 20, 30, 60, 90, 150, and 300. 1,000 random samples were used at each combination of n and quantile.

rates of the latter test becoming extremely liberal for the 0.95 and 0.99 quantile and $n <$ 90. Type I error rates for testing the intercept (H_0: $\beta_0 = 0$) with unweighted statistics deviated less from nominal rates compared to testing the slope (H_0: $\beta_1 = 0$) under similar heterogeneous error structures, providing reasonable Type I error rates for $n \geq 90$. The double permutation *F* test was not evaluated for this set of conditions but would be expected to provide similar improvements over the permutation *F* test as it did when error distributions were homogeneous.

Power for β_1 = 0.01, 0.05, 0.10, and 0.20 was simulated for $\gamma = 0.05$ for the unweighted rankscore tests because part of the motivation for using the rankscore tests was to avoid having to model error heterogeneity in applications. Clearly, slightly liberal Type I error rates for $\gamma = 0.05$ will inflate power estimates for the unweighted rankscore tests. Power for the unweighted *F* and *T* tests was similar, except for smaller *n* for 0.95 and 0.99 quantiles, where their Type I error rates had become excessively liberal or conservative, respectively (Fig. 2.7). Similar results were obtained for the normal and uniform error distributions (Appendices 2.7 and 2.8). Power for β_0 = 0.5, 1.0, 2.0, and 3.0 was slightly greater for the *T* compared to the *F* test (Fig. 2.7, Appendices 2.7 and 2.8), similar to homogeneous error distribution models.

4.4 Heterogenous Error Structure - Multiple Regression

The 6 parameter model, $y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + (1 + \gamma X_1)\varepsilon$, with $\gamma = 0.05$ was evaluated for the full model hypothesis H_0: $\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0$ for β_0 fixed at 36.0 and $\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0$ for Type I error rates, and with β_3 = 0.10, 0.15, 0.20, 0.25 for power. Type I error rates were well maintained by both tests until

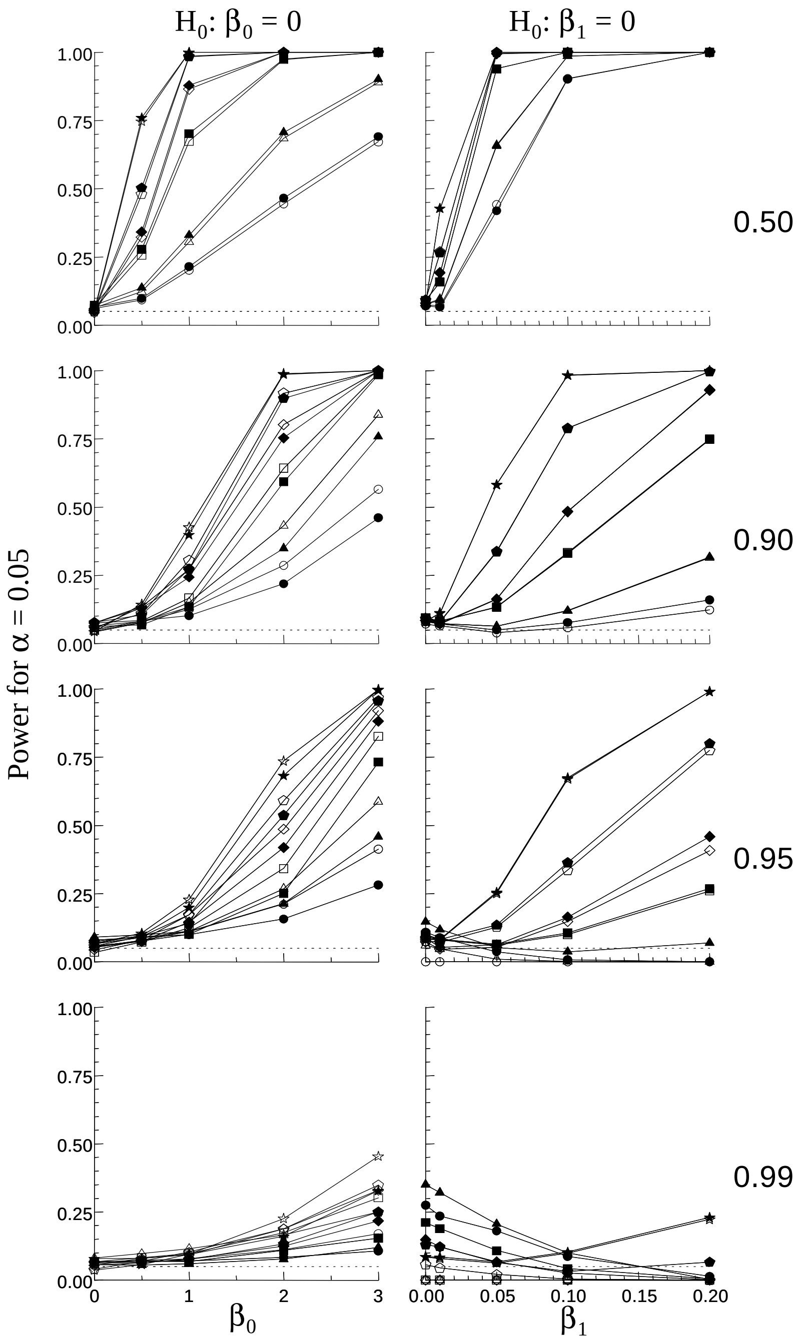


Figure 2.7. Estimated power for $\alpha = 0.05$ for the permutation F (solid) and Chi-square distributed T (open) rankscore tests; for heterogeneous lognormal error distributions; for H_0: $\beta_0 = 0$ and H_0: $\beta_1 = 0$ in the model $y = \beta_0 + \beta_1 X_1 + (1 + 0.05X_1)\varepsilon$; for $\beta_0 = 0.0$, 0.5, 1.0, 2.0, and 3.0 and for $\beta_1 = 0.0$, 0.01, 0.05, 0.10, and 0.20; for 0.50, 0.90, 0.95, and 0.99 quantiles; and for $n = 20$ (circle), 30 (triangle), 60 (square), 90 (diamond), 150 (pentagon), and 300 (star). Open symbols often are hidden behind solid symbols when equal. 1,000 random samples were used at each combination of effect size, n, and quantile.

$n \leq 30$ for the 0.95 quantile and $n \leq 150$ for the 0.99 quantile, where the *F* test became liberal and the *T* test became conservative (Fig. 2.8, Appendices 2.9 and 2.10). Power estimated with 1 of the 5 slope parameters (β_3) allowed to be nonzero was similar for the rankscore tests (Fig. 2.10). Power was low for the 0.95 quantile to nonexistent for the 0.99 quantile. Power for this and other conditions evaluated for the multiple regression models was only evaluated for the lognormal error distribution to reduce the amount of computing and reporting.

Type I error rates for subhypotheses involving continuous variables in the 6 parameter model with $\gamma = 0.05$ were evaluated for H_0: $\beta_3 = 0$ and H_0: $\beta_3 = \beta_5 = 0$ with $\beta_0 = 36.0$, $\beta_1 = 0.10$, $\beta_2 = -0.005$, $\beta_4 = 2.0$, and $\beta_3 = \beta_5 = 0.0$. The permutation *F* test maintained Type I errors well across all sample sizes and quantiles for H_0: $\beta_3 = 0$, whereas the *T* test became excessively conservative for smaller *n* for 0.95 and 0.99 quantiles (Fig. 2.8). Type I error rates were slightly more liberal for the H_0: $\beta_3 = \beta_5 = 0$ (Fig. 2.9) compared to the H_0: $\beta_3 = 0$ (Fig. 2.8) for lognormal as well as normal and uniform error distributions (Appendices 2.9 - 2.12). Again, the permutation *F* test maintained Type I error rates better for smaller *n* and 0.95 and 0.99 quantiles compared to the *T* test, which became excessively conservative. Power for H_0: $\beta_3 = 0$ was estimated with $\beta_3 = 0.10$, 0.15, 0.20, and 0.25 for the lognormal error distribution. Power was similar for the tests and became exceedingly low to nonexistent for 0.90 - 0.99 quantiles (Fig. 2.10).

Subhypotheses involving categorical predictors in the 6 parameter model were evaluated for H_0: $\beta_4 = 0$ and H_0: $\beta_4 = \beta_5 = 0$ with $\beta_0 = 36.0$, $\beta_1 = 0.10$, $\beta_2 = -0.005$,

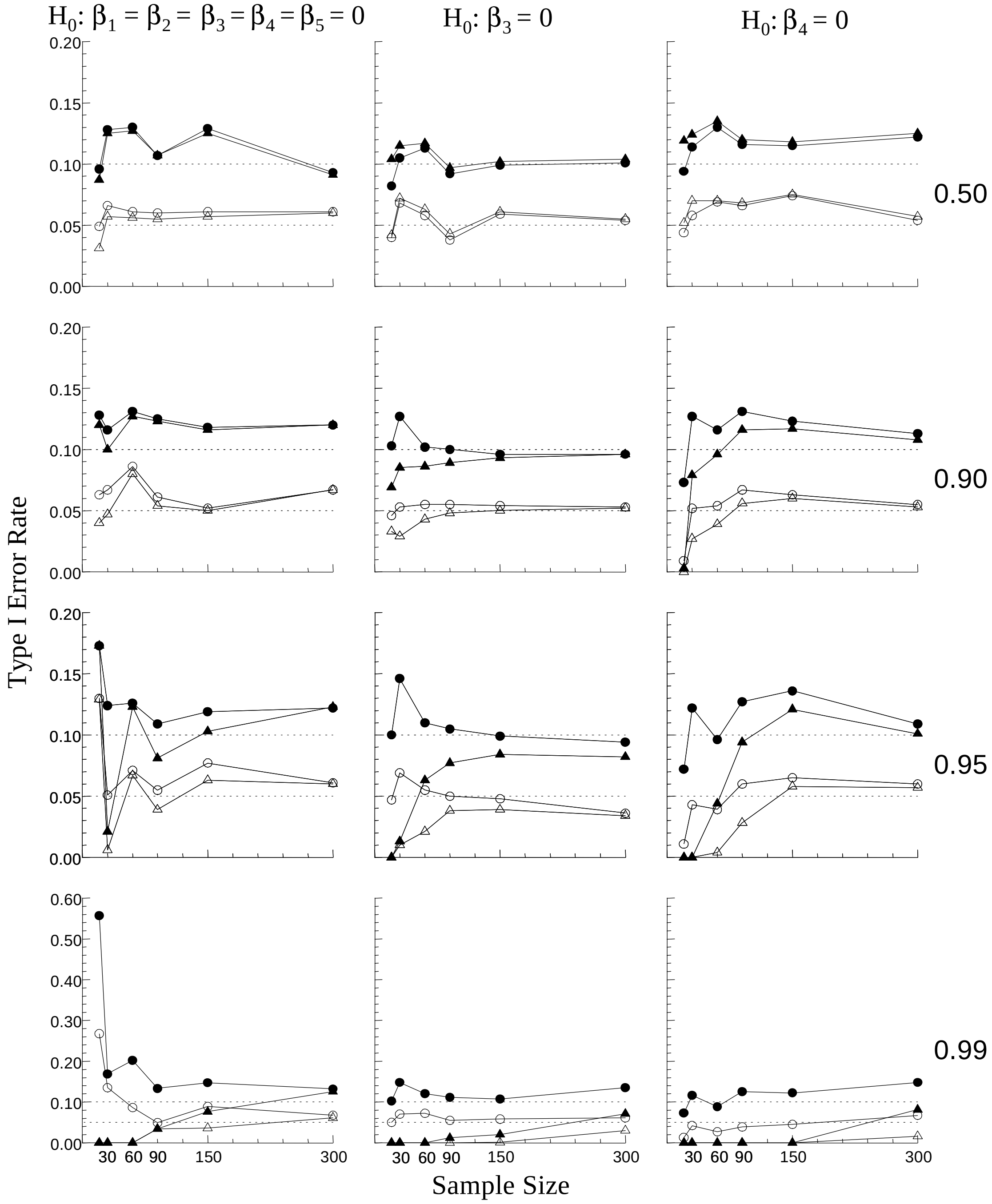


Figure 2.8. Estimated type I error rates for $\alpha = 0.05$ (open) and 0.10 (solid); for the permutation F (circles) and Chi-square distributed T (triangles) rankscore tests for H_0: $\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0$, H_0: $\beta_3 = 0$, and H_0: $\beta_4 = 0$; for heterogeneous lognormal error distributions with $\gamma = 0.05$ in the model $y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + (1 + \gamma X_1)\varepsilon$; for 0.50, 0.90, 0.95, and 0.99 quantiles; and for n = 20, 30, 60, 90, 150, and 300. 1,000 random samples were used at each combination of H_0, n, and quantile.

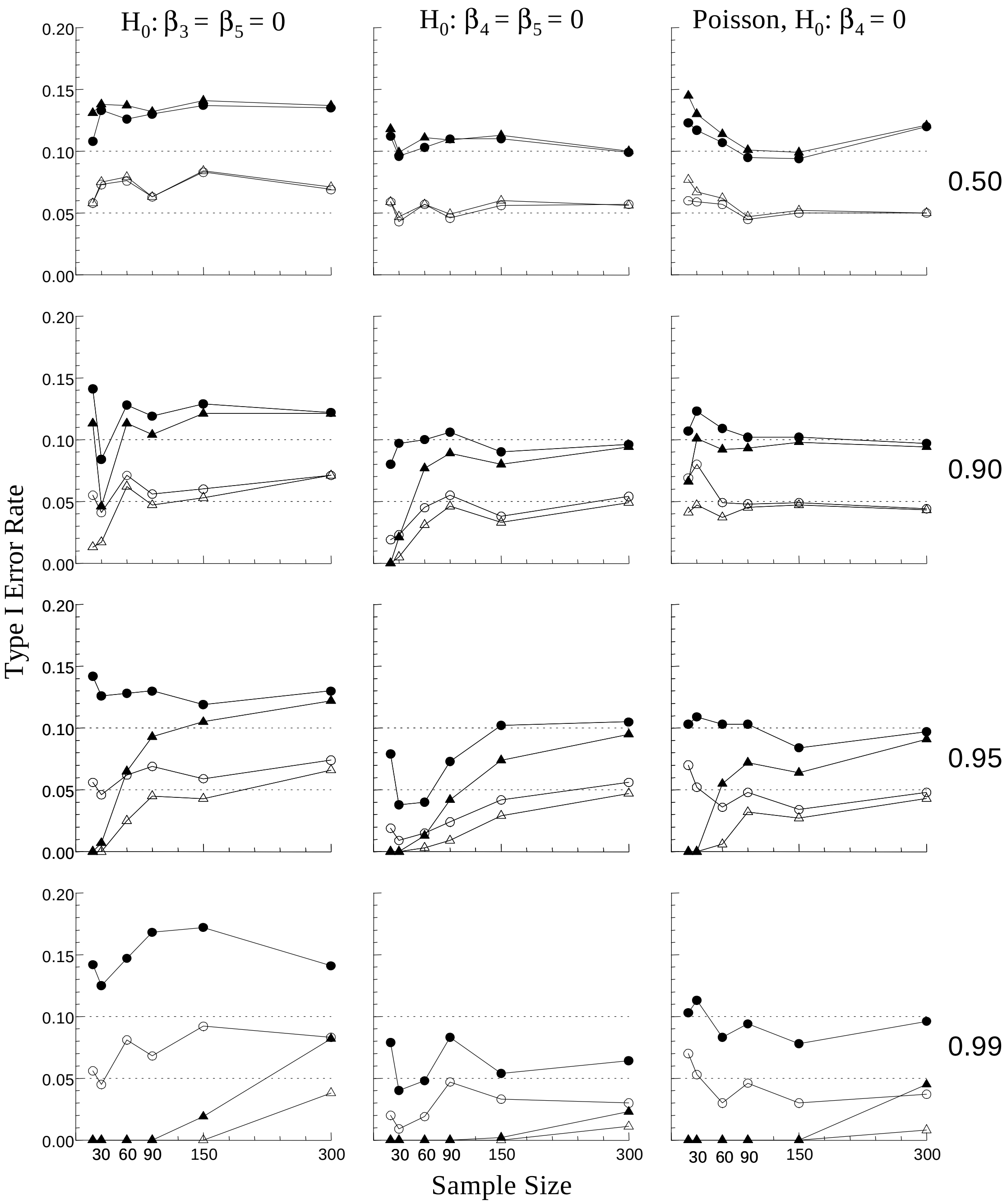


Figure 2.9. Estimated type I error rates for $\alpha = 0.05$ (open) and 0.10 (solid); for the permutation F (circles) and Chi-square distributed T (triangles) rankscore tests for H_0: $\beta_3 = \beta_5 = 0$ and H_0: $\beta_4 = \beta_5 = 0$ for heterogeneous lognormal error distributions with $\gamma = 0.05$ in the model $y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + (1 + \gamma X_1)\varepsilon$; and for H_0: $\beta_4 = 0$ in the model $y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5$ where y has a Poisson distribution; for 0.50, 0.90, 0.95, and 0.99 quantiles; and for n = 20, 30, 60, 90, 150, and 300. 1,000 random samples were used at each combination of H_0, n, and quantile.

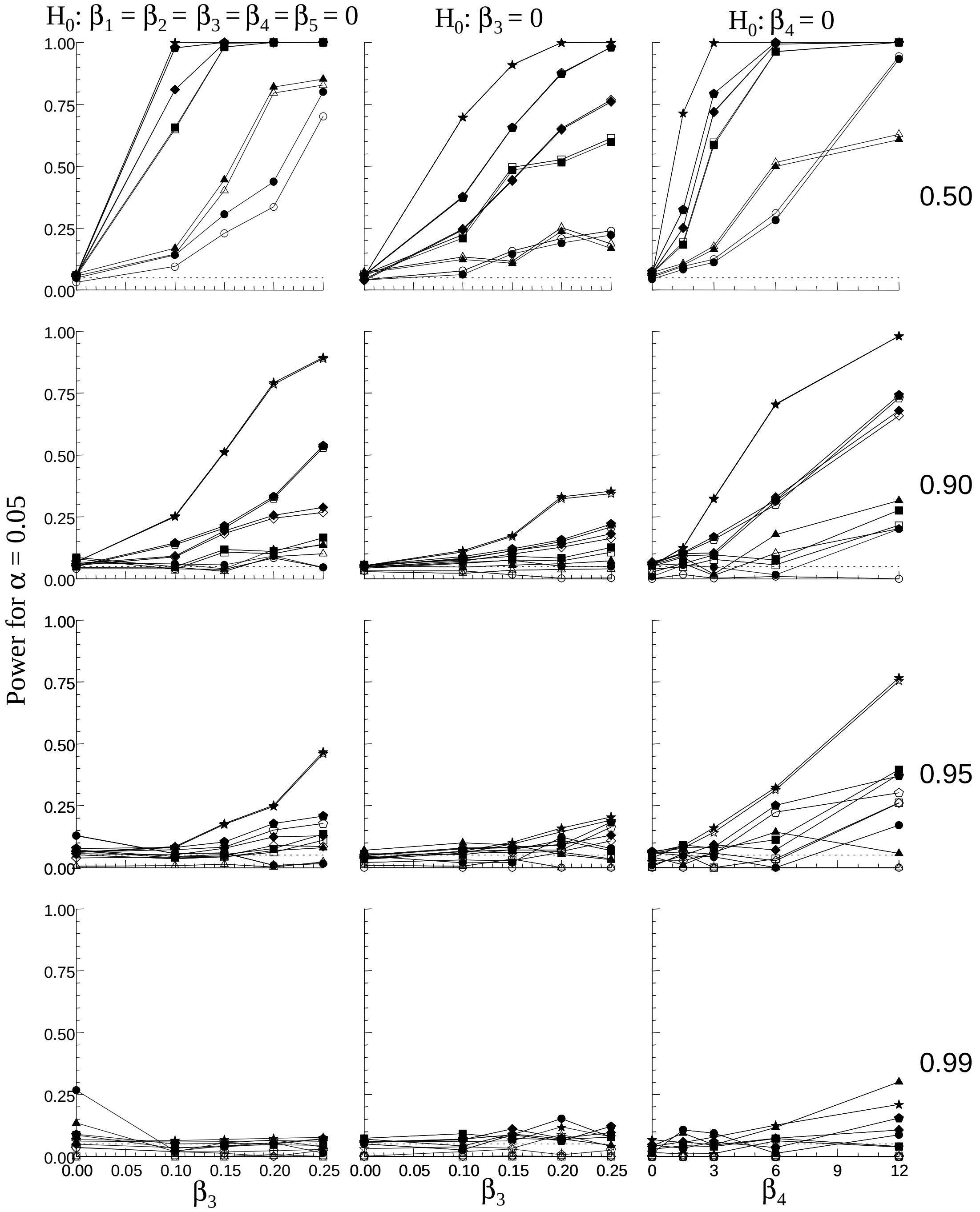


Figure 2.10. Estimated power for $\alpha = 0.05$ for the permutation F (solid) and Chi-square distributed T (open) rankscore tests for H_0: $\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0$, H_0: $\beta_3 = 0$, and H_0: $\beta_4 = 0$; for heterogeneous lognormal error distributions with $\gamma = 0.05$ in the model $y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + (1 + \gamma X_1)\varepsilon$; for 0.50, 0.90, 0.95, and 0.99 quantiles; and for n = 20 (circle), 30 (triangle), 60 (square), 90 (diamond),150 (pentagon), and 300 (star). Open symbols often are hidden behind solid symbols when equal. 1,000 random samples were used at each combination of H_0, n, and quantile.

$\beta_3 = 0.05$, and $\beta_4 = \beta_5 = 0.0$. The T test became excessively conservative for $n \leq 90$ for the 0.95 quantile and for $n \leq 150$ for the 0.99 quantile compared to the permutation F test (Figs. 2.8 and 2.9). Type I error rates for H_0: $\beta_4 = 0$ (Fig. 2.8) were slightly more liberal than for H_0: $\beta_4 = \beta_5 = 0$ (Fig. 2.9) for both tests for lognormal as well as normal and uniform error distributions (Appendices 2.9 - 2.12). Power was evaluated for the subhypothesis H_0: $\beta_4 = 0$ for $\beta_4 = 1.5$, 3.0, 6.0, and 12.0 and the lognormal error distribution. Estimates of power were similar for the tests with a slight advantage for the permutation F test for smaller n and the 0.95 and 0.99 quantiles (Fig.2.10).

The H_0: $\beta_4 = 0$ also was evaluated for a variant of this 6 parameter model where $\beta_0 = 3.0$, $\beta_1 = 0.10$, $\beta_2 = -0.0005$, $\beta_3 = 0.05$, and $\beta_4 = \beta_5 = 0.0$ and y having a Poisson distribution with mean and variance specified by the regression function. As elsewhere, the permutation F test maintained better error rates for small n for the 0.95 and 0.99 quantiles than the T test (Fig. 2.9). For this model, there was no evidence that the tied integer values associated with the Poisson distribution caused any unusual problems with the rankscore tests.

The F distribution approximation of the $F_{q,\, n-p}$ rankscore statistic maintained Type I error rates well under similar sample sizes and quantiles where the Chi-square distributional approximation of the T rankscore statistic worked well when testing subhypotheses in multiple regression models. However, probabilities for the $F_{q,\, n-p}$ statistic and those provided by the permutation approximation of the F_o statistic were closer to nominal error rates for small n and more extreme quantiles than those for the Chi-square distributional approximation of the T statistic. An example for H_0: $\beta_4 = \beta_5 =$

0 for the lognormal error distribution and the 0.95 quantile is in Figure 2.11.

5. Example Application

I constructed confidence intervals for regression quantile estimates of Lahontan cutthroat trout *Oncorhynchus clarki henshawi* density (trout m^{-1}) as a function of stream channel morphology (width:depth ratio) for 13 small streams in Nevada sampled over 7 years (Dunham et al. 2002). Width:depth ratio is a measure that integrates stream channel characteristics thought to be related to small stream integrity and, thus, fish populations and is easily measured for assessing fish habitat conditions and land use impacts over large regions. Lahontan cutthroat trout are a threatened species of special interest to federal land management agencies.

Here I considered the nonlinear model $y = \exp(\beta_0 + \beta_1 X_1 + \varepsilon)$, where y is trout m^{-1} and X_1 is width:depth ratio, for n = 71 observations of streams for 1993 to 1999 (Dunham et al. 2002). The model was estimated in the linear form $\ln y = \beta_0 + \beta_1 X_1 + \varepsilon$ and estimates for selected regression quantiles were plotted by exponentiating to back transform to the nonlinear form (Fig. 2.12). Estimates for all quantiles were plotted as a step function with 90% confidence intervals for 19 quantiles between 0.05 and 0.95 by increments of 0.05 (Fig. 2.13). Interval endpoints were estimated from a linear interpolation between hypothesized parameter values that had T test statistics that bracketed the standard normal test statistic = 1.645 associated with α = 0.10 (Koenker 1994) as was done by Dunham et al. (2002). Here I also provide confidence intervals for estimates of β_0 that were not provided by Dunham et al. (2002).

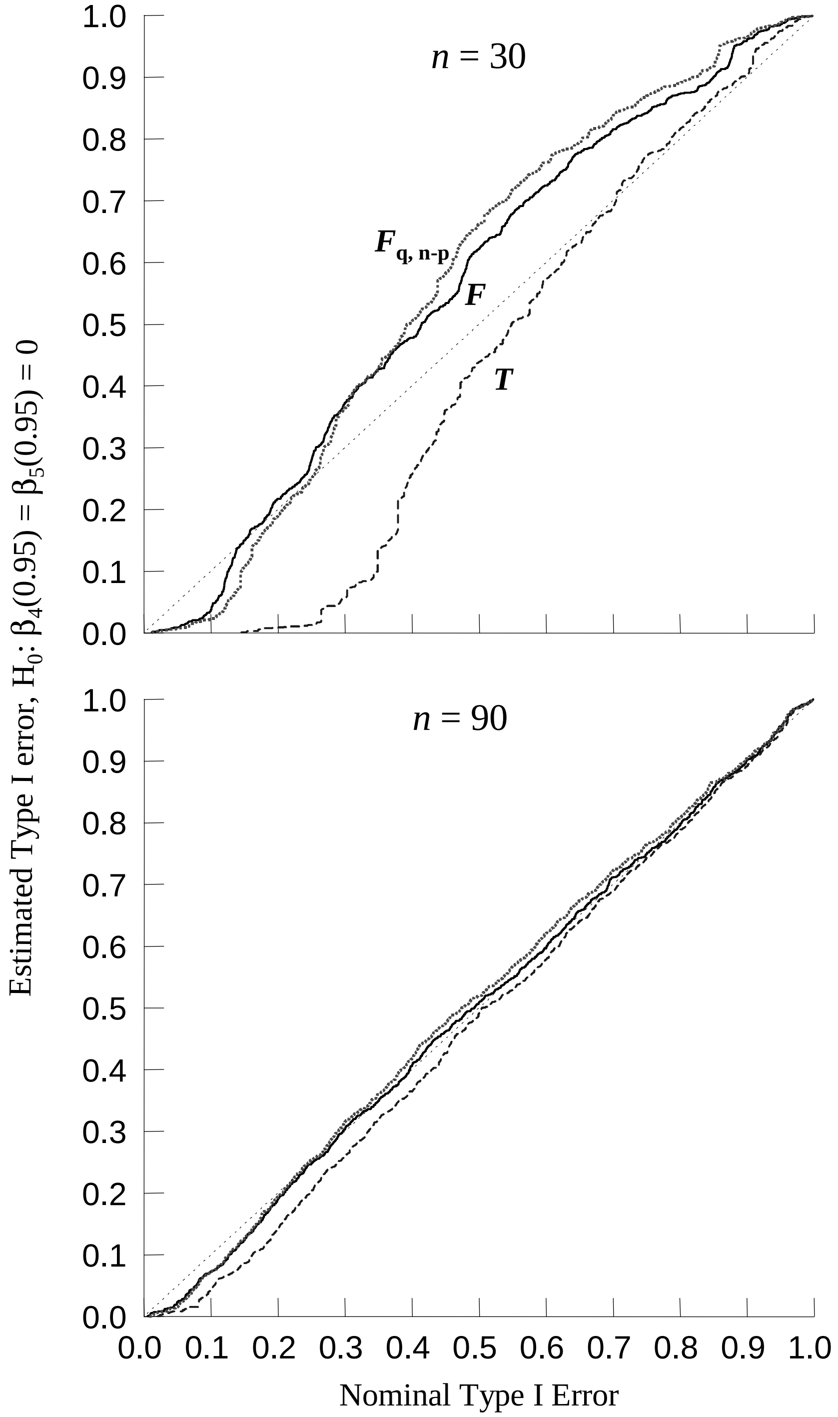


Figure 2.11. Cumulative distributions of 1,000 estimated errors for permutation approximation of the *F* (solid line), $F_{q,\,n-p}$ distribution approximation of (square dot) the *F*, and Chi-square distribution approximation of *T* (dashes) rankscore tests for H_0: $\beta_4 = \beta_5 = 0$ for the 0.95 quantile, for n = 30 and 90, for the lognormal error distribution in the model $y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + (1 + \gamma X_1)\varepsilon$.

Confidence intervals also were constructed based on inverting the permutation *F* test for the same quantiles (Fig. 2.13). The possible boundary values for the estimated confidence interval endpoints were obtained from the linear programming implementation used to construct intervals by inverting the *T* test statistic (Koenker 1994). These values were then used as hypothesized parameter values of $\xi(\tau)$ in the transformation $\mathbf{y} - \mathbf{X_2}\xi(\tau)$ to test the H_0: $\boldsymbol{\beta}_2(\tau) = \xi(\tau)$ with (6), where $\boldsymbol{\beta}_2$ was either β_0 or β_1 depending on the parameter being tested. We used $m + 1 = 100{,}000$ permutations to compute probabilities for the *F* tests associated with confidence interval endpoints. Similar to the *T* test inversion approach, the *F* test inversion approach had confidence interval endpoints that were discontinuous in probabilities. I used a linear interpolation based on the *P*-values to estimate the endpoints rather than the more conservative approach of using the closest estimated confidence interval endpoint with $P \leq \alpha$. This had a similar effect to the linear interpolation for the *T* test. For example, the hypothesized parameter values that bracketed the lower 90% confidence interval endpoint for the 0.90 quantile for β_1 were $\xi(0.90) = -0.03374$ with $P = 0.0396$ and $\xi(0.90) = -0.03346$ with $P = 0.2980$. No value between these parameter values yielded different rankscore test statistics. The linear interpolated interval was computed as $-0.03374 + |-0.03346 - -0.03374| \times ((0.1000 - 0.0396)/(0.2980 - 0.0396)) = -0.03367$. The *T* test inversion approach has standard normal test statistics of 2.047 and 1.241 associated with $\xi(0.90) = -0.03374$ and $\xi(0.90) = -0.03346$, respectively. Linear interpolation was used to obtain the estimated endpoint (-0.03361) for the standard normal test statistic = 1.645 associated with $\alpha = 0.10$ (Koenker 1994). There was little

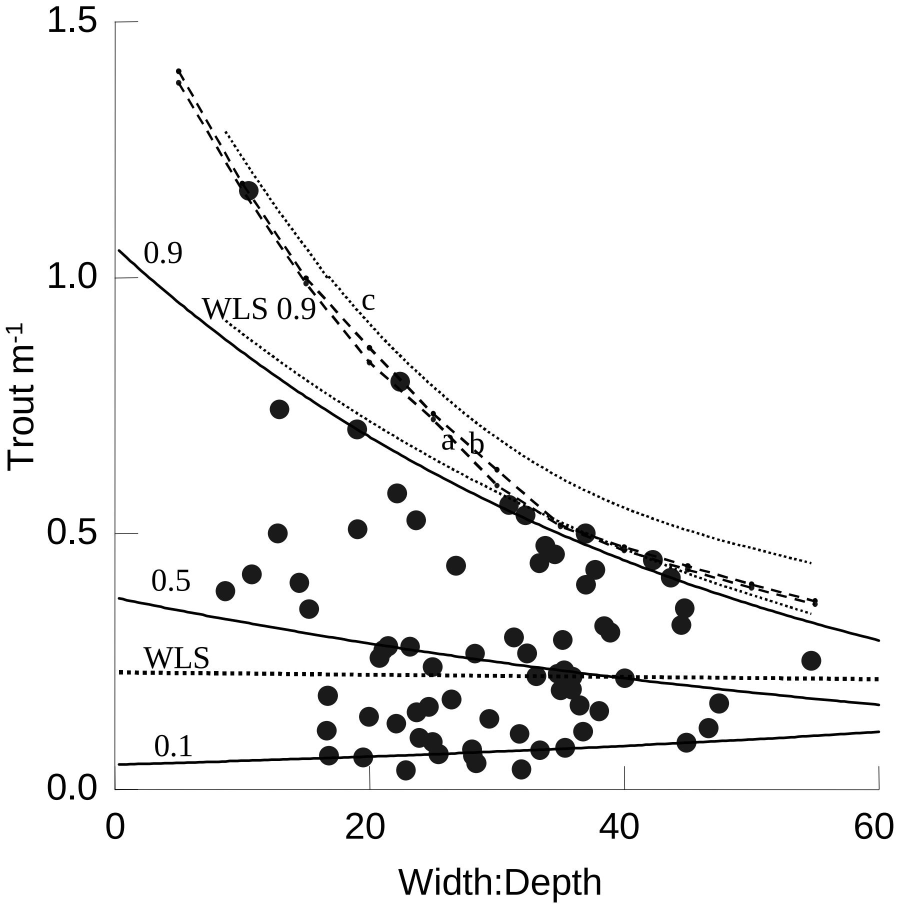


Figure 2.12. Lahontan cutthroat trout m^{-1} and width:depth ratios for 13 small streams sampled 1993-1999 (n = 71); exponentiated estimates for 0.90, 0.50, and 0.10 regression quantiles (solid lines) for the model $\ln y = \beta_0 + \beta_1 X_1 + \varepsilon$; and exponentiated weighted least squares (WLS) estimate of mean and 0.90 percentile (WLS 0.9) estimate for the model $\ln yw = (\beta_0 + \beta_1 X_1 + (\gamma_0 - \gamma_1 X_1)\varepsilon)w$, $w = (1.310 - 0.017X_1)^{-1}$ (dotted lines). Dashed lines are nonsimultaneous (a) and simultaneous (b) 1-sided upper 90% confidence intervals for 0.90 regression quantile for selected width:depth ratios between 5 and 55. Upper dotted line (c) is nonsimultaneous 1-sided upper 90% confidence interval for 0.90 percentile estimate based on the weighted least squares model.

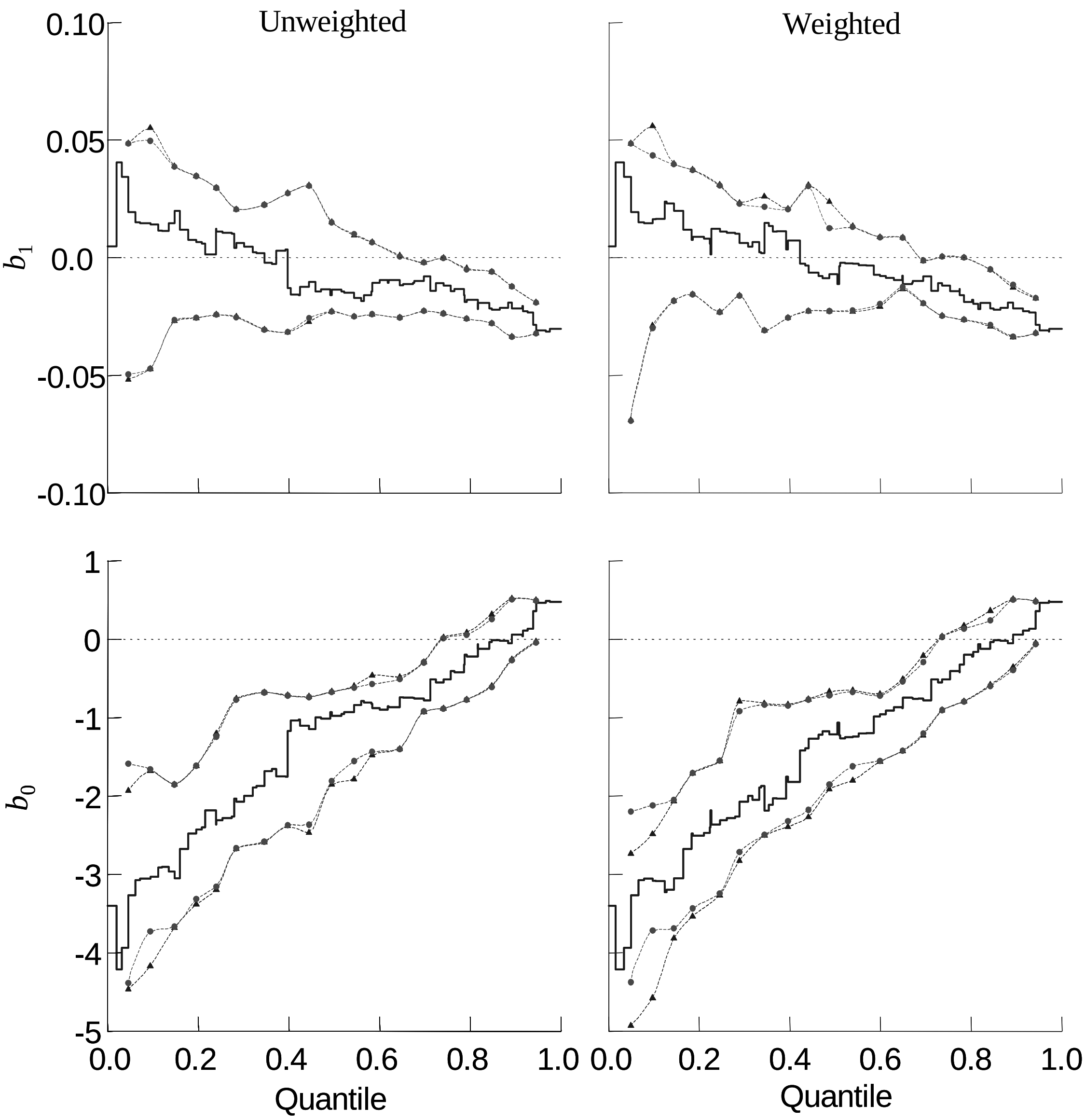


Figure 2.13. Solid lines are step functions for estimates of β_0 and β_1 by quantiles [0, 1] in the unweighted model ln $y = \beta_0 + \beta_1 X_1 + \varepsilon$ and in the weighted model $(\ln y)w = (\beta_0 + \beta_1 X_1 + (\gamma_0 - \gamma_1 X_1)\varepsilon)w$, $w = (1.310 - 0.017X_1)^{-1}$, for n =71 observations of Lahontan cutthroat trout m^{-1} and width:depth ratios. Pointwise 90% confidence intervals based on inverting the T rankscore test (triangles) and inverting the permutation F rankscore test (cirlcles) were constructed with linear interpolation between estimated endpoints.

difference in the estimated 90% confidence intervals across τ = [0.05, 0.95] for the permutation F and T test inversion approaches, with slightly narrower intervals for the permutation F inversion approach for some quantiles (Fig. 2.13). Pushing interval estimation for β_1 to a slightly more extreme quantile, $\tau = 0.98$, the permutation F test based interval (-0.0313, -0.0299) was a third of the length of the T test based interval (-0.0324, -0.0293), although both intervals were very short and perhaps stretched the bounds of reliability.

Weighted regression quantile estimates and associated intervals based on test inversion were constructed by using the unweighted estimates (Fig. 2.13) as a basis for developing a weighting function. The decrease in estimates of β_1 mirrors the increase in estimates of β_0 with increasing τ, suggesting a linear location-scale model with error variation of the form $\gamma_0 - \gamma_1 X_1$ as a reasonable approximation. The weight function was estimated by the average pairwise difference between the 76 regression quantile estimates for $b_0(\tau)$ to estimate γ_0 and for $b_1(\tau)$ to estimate γ_1. Multiresponse permutation procedure routines were used for computing the average pairwise differences (Mielke and Berry 2001). The estimated standard deviation function was $1.310 - 0.017X_1$, and its reciprocal provided weights for the weighted regression quantile estimate (2), which was implemented by multiplying all variables in the model by the weights and then using the regression quantile estimator (1). The F and T tests produced linear interpolated 90% confidence intervals that differed most for weighted estimates of β_0 for the lower quantiles (Fig. 2.13). The overall pattern and width of intervals for the weighted estimates were not greatly different from their unweighted counterparts, which is consistent with the rather weak (<1 standard deviation change) pattern of heterogeneity across width:depth ratios. Both weighted and unweighted confidence bands supported an interpretation that increasing stream width:depth ratios from 15 to 45 decreased the highest 20% of trout densities ($\tau \geq 0.80$) by 11 to 64% [exp(-0.004 × 30) = 0.887 and exp(-0.034 × 30) = 0.361].

A 1-sided upper 90% confidence band for the 0.90 quantile that was not simultaneous in X_1 was estimated for 11 equally spaced width:depth values between 5 and 55 corresponding to the range of ratios in the sample (Fig. 2.12). This was done by forming confidence intervals for β_0 with a 2-sided $\alpha = 0.20$ after shifting the width:depth ratios by the 11 selected values. For example, shifting the width:depth ratios by subtracting 20 implies that the interval constructed for β_0 on the transformed data was now an interval for width:depth ratio = 20 rather than for width:depth ratio = 0. Obviously, more values of width:depth ratio could have been used to obtain a smoother band. For comparison, a 90th percentile line based on a weighted least squares regression of the log transformed trout densities and corresponding nonsimultaneous 1-sided upper 90% confidence intervals were estimated based on Vardeman (1992) and Gerow and Bilen (1999). Both the quantile regression and weighted least squares intervals are interpreted as upper tolerance intervals for an individual value of width:depth (Vardeman 1992), but the latter estimates assumed a normal distribution for the log transformed data, resulting in slightly wider intervals (Fig. 2.12). A lower confidence interval (e.g., for 0.10 quantile) was of little interest with this data as it was effectively 0 for all width:depth ratios.

Simultaneous intervals in X_1 for the 0.90 quantile regression were estimated by emulating computations for the Working-Hotelling procedure for simultaneous confidence bands (Neter et al. 1996:156-157). The simultaneous intervals were slightly wider than the nonsimultaneous quantile rankscore intervals but were narrower than the nonsimultaneous intervals based on the weighted least squares estimates (Fig. 2.12). The Working-Hotelling procedure used $(2 \times F(0.80, 2, 69))^{0.5} = 1.815$ as a multiplier for the standard error of a predicted y at a specified X, implying that any individual interval required an $\alpha = 0.0738$ for a simultaneous 2-tailed $\alpha = 0.20$. The simultaneous confidence band in Figure 2.12 is interpreted as an upper 90% tolerance band for 90% of future observations of trout densities. In repeated random sampling we would expect 90% of samples to have 90% of trout densities within the interval estimates for all width:depth ratios. Although I used the T rankscore test inversion approach for constructing the confidence bands on the 0.90 quantile, this procedure also could have been done with the permutation F test inversion procedure. A simultaneous confidence band also could be constructed based on the weighted least squares estimates (Turner and Bowden 1977, Gerow and Bilen 1999) but would be even wider than the nonsimultaneous band.

6. Discussion

The permutation F rankscore test maintained Type I errors better and had more power than the Chi-square T rankscore test for model combinations of small samples, more extreme quantiles, and more parameters. The permutation test maintained Type I errors better at $20 \leq n \leq 30$ for 0.95 and $20 \leq n \leq 90$ for 0.99 quantiles for 2 parameter models and at $30 \leq n \leq 90$ for 0.95 and $150 \leq n < 300$ for 0.99 quantiles for 6 parameter models, depending on the number of parameters being tested. This was true regardless of the error distribution. My example application with the Lahontan cutthroat trout data suggested that these differences may not always be of sufficient magnitude to affect the interpretation of an analysis when quantiles used are not too extreme (e.g., $0.05 \leq \tau \leq 0.95$). When estimating models for more extreme quantiles (e.g., $\tau = 0.99$), fairly large samples ($n > 300$) will be required for models with more than just a few parameters to ensure reliable confidence intervals by either test. Power to detect the alternative hypothesis was low for more extreme quantiles (0.95 and 0.99) in the low density tails of the lognormal and normal error distributions. This results in wider estimated confidence intervals based on test inversion. This was especially problematic for testing subhypotheses in models with more parameters. The F distribution approximation of the $F_{q,\, n-p}$ form of the rankscore statistic offered some advantages over the Chi-square distribution approximation of the T rankscore statistic at small n and more extreme quantiles when testing subhypotheses in multiple regression models. But there was greater improvement by going to the permutation approximation of the F rankscore statistic, at least for parameters other than the intercept.

The double permutation scheme (Legendre and Desdevises In Press) provided better Type I errors for the F test when null models were forced through the origin as when testing the intercept. However, additional refinements of the double permutation scheme need to be investigated to see whether this approach can be made to work as well for more extreme quantiles (e.g., 0.99) and small samples. This will be especially important in applications of permutation tests with weighted models.

The rankscore tests were not immune to the effects of heterogeneity, although this was a more serious performance issue for the simple 2 parameter regression models than for the 6 parameter multiple regression models. Some adjustment for error heterogeneity will often be desirable for the regression quantile rankscore tests and confidence intervals. My simulation results suggested that when there was variation across an independent variable >2.5 standard deviations, tests and confidence interval estimates might benefit from using weighted estimates and rankscore tests. I used a simple pairwise difference approach based on the initial unweighted estimates for estimating weights in the example application. Other approaches for estimating weights include regressing absolute values of residuals from an unweighted fit of the 0.5 quantile on the independent variables for linear location-scale models (Zhou and Portnoy 1998) and the sparsity estimation approach for more general heteroscedastic models (Koenker and Machado 1999).

One of the potential benefits of analyzing data with regression quantiles is to focus attention on the utility of prediction and tolerance intervals in the linear model (Vardeman 1992). My simulations established the validity of the quantile rankscore tests for constructing confidence intervals for β_0, and, therefore, by implication for other values of $\mathbf{X} = \mathbf{x}$. Inverting tests on appropriate regression quantile estimates allows construction of prediction and tolerance intervals without assuming a specific form of the error distribution. Zhou and Portnoy (1998) provided alternative order statistic based approaches to constructing such intervals with regression quantiles. The quantile regression based tolerance intervals estimated in my example application were slightly narrower than comparable intervals based on weighted least squares estimates that assumed a normal error distribution. Parametric distributional approaches for setting prediction and tolerance intervals should provide narrower intervals only when the distributional assumptions are well founded. This will not be common in most ecological and biological applications. Recall that the assumed parametric error distributional form is of less consequence when estimating parameters and intervals associated with the conditional mean than it is when trying to estimate parameters associated with other parts of the probability distribution, as is required for constructing prediction and tolerance intervals.

Acknowledgments

J. D. Richards provided programming support for the simulations. M. J. Anderson, P. Good, R. Koenker, and J. E. Roelle reviewed earlier drafts of the manuscript. J. B. Dunham provided the Lahontan cutthroat trout data.

Literature Cited

Allen, A. W., B. S. Cade, and M. W. Vandever. 2001. Effects of emergency haying on vegetative characteristics within selected conservation reserve program fields in the northern Great Plains. J. Soil and Water Conservation **56**: 120-125.

Anderson, M. J., and P. Legendre. 1999. An empirical comparison of permutation methods for tests of partial regression coefficients in a linear model. Journal

Statistical Computation and Simulation **62**:271-303.

Anderson, M. J., and J. Robinson. 2001. Permutation tests for linear models. Australian New Zealand Journal of Statistics **43**:75-88.

Barrodale, I., and F. D. K. Roberts. 1974. Algorithm 478: Solution of an overdetermined system of equations in the l_1 norm. Communications of the Association for Computing Machinery **17**:319-320.

Buchinsky, M. 1991. The theory and practice of quantile regression. Ph.D dissertation, Harvard University, Cambridge Mass. USA. 211pp.

Buchinsky, M. 1995. Quantile regression, Box-Cox transformation model, and the U.S. wage structure, 1963-1987. Journal of Econometrics **65**: 109-154.

Cade, B. S., and Q. Guo. 2000. Estimating effects of constraints on plant performance with regression quantiles. Oikos **91**: 245-254.

Cade, B. S., and J. D. Richards. 1996. Permutation tests for least absolute deviation regression. Biometrics **52**:886-902.

Cade, B. S., J. W. Terrell, and R. L. Schroeder. 1999. Estimating effects of limiting factors with regression quantiles. Ecology **80**:311-323.

Chernozhukov, V., and L. Umantsev. 2001. Conditional value-at-risk: Aspects of modeling and estimation. Empirical Economics **26**: 271-292.

Commenges, D. In Press. Transformations which preserve exchangeability and application to permutation tests. Journal of Nonparametric Statistics.

Dunham, J. B., B. S. Cade, and J. W. Terrell. 2002. Influences of spatial and temporal variation on fish-habitat relationships defined by regression quantiles.

Transactions of the American Fisheries Society **131**: 86-98.

Eastwood, P. D., G. J. Meaden, and A. Grioche. 2001. Modeling spatial variations in spawning habitat suitability for the sole *Solea solea* using regression quantiles and GIS procedures. Marine Ecology Progress Series **224**: 251-266.

Freedman, D., and D. Lane. 1983. A nonstochastic interpretation of reported significance levels. Journal of Business and Economic Statistics **1**:292-298.

Gerow, K., and C. Bilen. 1999. Confidence intervals for percentiles: An application to estimation of potential maximum biomass of trout in Wyoming streams. North American Journal of Fisheries Management **19**:149-151.

Gutenbrunner, C., and J. Jurečková. 1992. Regression rank scores and regression quantiles. Annals of Statisitics **20**: 305-330.

Gutenbrunner, C., J. Jurečková, R. Koenker, and S. Portnoy. 1993. Tests of linear hypotheses based on regression rank scores. Nonparametric Statistics **2**:307-331.

Haire, S. L., C. E. Bock, B. S. Cade, and B. C. Bennett. 2000. The role of landscape and habitat characteristics in limiting abundance of grassland nesting songbirds in an urban open space. Landscape and Urban Planning **48**(1-2):65-82.

Huston, M. A. 2002. Introductory essay: Critical issues for improving predictions. Pages 7-21 *in* J. M. Scott, P. J. Heglund, and M. L. Morrison, editors. Predicting species occurrences: issues of accuracy and scale. Island Press, Covelo, California, USA.

Kaiser, M. S., P. L. Speckman, and J. R. Jones. 1994. Statistical models for limiting nutrient relations in inland waters. Journal of the American Statistical Association **89**:410-423.

Kennedy, P. E., and B. S. Cade. 1996. Randomization tests for multiple regression. Communications in Statistics - Simulation and Computation **25**: 923-936.

Koenker, R. 1987. A comparison of asymptotic testing methods for l_1-regression. Pages 287-295 *in* Y. Dodge, editor. Statistical data analysis based on the L_1-norm and related methods. Elsevier Science Publishers B. V. (North-Holland), Amsterdam.

Koenker, R. 1994. Confidence intervals for regression quantiles. *Pages* 349-359 *in* P. Mandl and M. Hušková, editors. Asymptotic statistics: Proceedings of the 5th Prague Symposium. Physica-Verlag

Koenker, R. 1997. Rank tests for linear models. Pages 175-199 *in* G. S. Maddala and C. R. Rao, editors. Handbook of Statistics, vol 15.

Koenker, R., and G. Bassett. 1978. Regression quantiles. Econometrica **46**:33-50.

Koenker, R., and G. Bassett. 1982. Robust tests for heteroscedasticity based on regression quantiles. Econometrica **50**: 43-61.

Koenker, R. and V. d'Orey. 1987. Computing regression quantiles. Applied Statistics **36**:383-393.

Koenker, R. and V. d'Orey. 1994. A remark on algorithm AS229: computing dual regression quantiles and regression rank scores. Applied Statistics **43**:410-414.

Koenker, R., and O. Geling. 2001. Reappraising medfly longevity: a quantile regression survival analysis. Journal American Statistical Association **96**: 458-

468.

Koenker, R., and K. F. Hallock. 2002. Quantile regression: An introduction. Journal Economic Perspectives, In press.

Koenker, R., and J. A. F. Machado. 1999. Goodness of fit and related inference processes for quantile regression. Journal of the American Statistical Association **94**:1296-1310.

Koenker, R., and S. Portnoy. 1996. Quantile regression. University of Illinois at Urbana-Champaign, College of Commerce and Business Administration, Office of Research Working Paper 97-0100. 77pp.

Koenker, R., and Q. Zhao. 1994. *L*-estimation for linear heteroscedastic models. Nonparametric Statistics **3**:223-235.

Legendre, P. 2000. Comparison of permutation methods for the partial correlation and partial Mantel tests. Journal of Statistical Computation and Simulation **67**:37-73.

Legendre, P., and Y. Desdevises. (In press). Independent contrasts and regression through the origin. Systematic Biology.

Mansouri, H. 1999. Aligned rank transform tests in linear models. Journal of Statistical Planning and Inference **79**:141-155.

Mielke, P. W., Jr., and K. J. Berry. 2001. Permutation Methods: A Distance Function Approach. Springer-Verlag, New York. 352pp.

Neter, J., M. H. Kutner, C. J. Nachtsheim, and W. Wasserman. 1996. Applied linear statistical models. Irwin, Chicago, Illinois, USA. 1408pp.

Randles, R. H. 1984. On tests applied to residuals. Journal of the American Statistical Association **79**:349-354.

Rosenbaum, P. R. 1991. Discussing hidden bias in observational studies. Annals of Internal Medicine **115**:901-905.

Rosenbaum, P. R. 1995. Quantiles in nonrandom samples and observational studies. Journal of the American Statistical Association **90**:1424-1431.

Scharf, F. S., F. Juanes, and M. Sutherland. 1998. Inferring ecological relationships from the edges of scatter diagrams: comparison of regression techniques. Ecology **79**:448-460.

Terrell, J. W., B. S. Cade, J. Carpenter, and J. M. Thompson. 1996. Modeling stream fish habitat limitations from wedged-shaped patterns of variation in standing stock. Transactions of the American Fisheries Society **125**:104-117.

Theil, H. 1965. The analysis of disturbances in regression analysis. Journal of the American Statistical Association **60**:1067-1079.

Thomson, J. D., G. Weiblen, B. A. Thomson, S. Alfaro, and P. Legendre. 1996. Untangling multiple factors in spatial distributions: Lilies, gophers, and rocks. Ecology **77**:1698-1715.

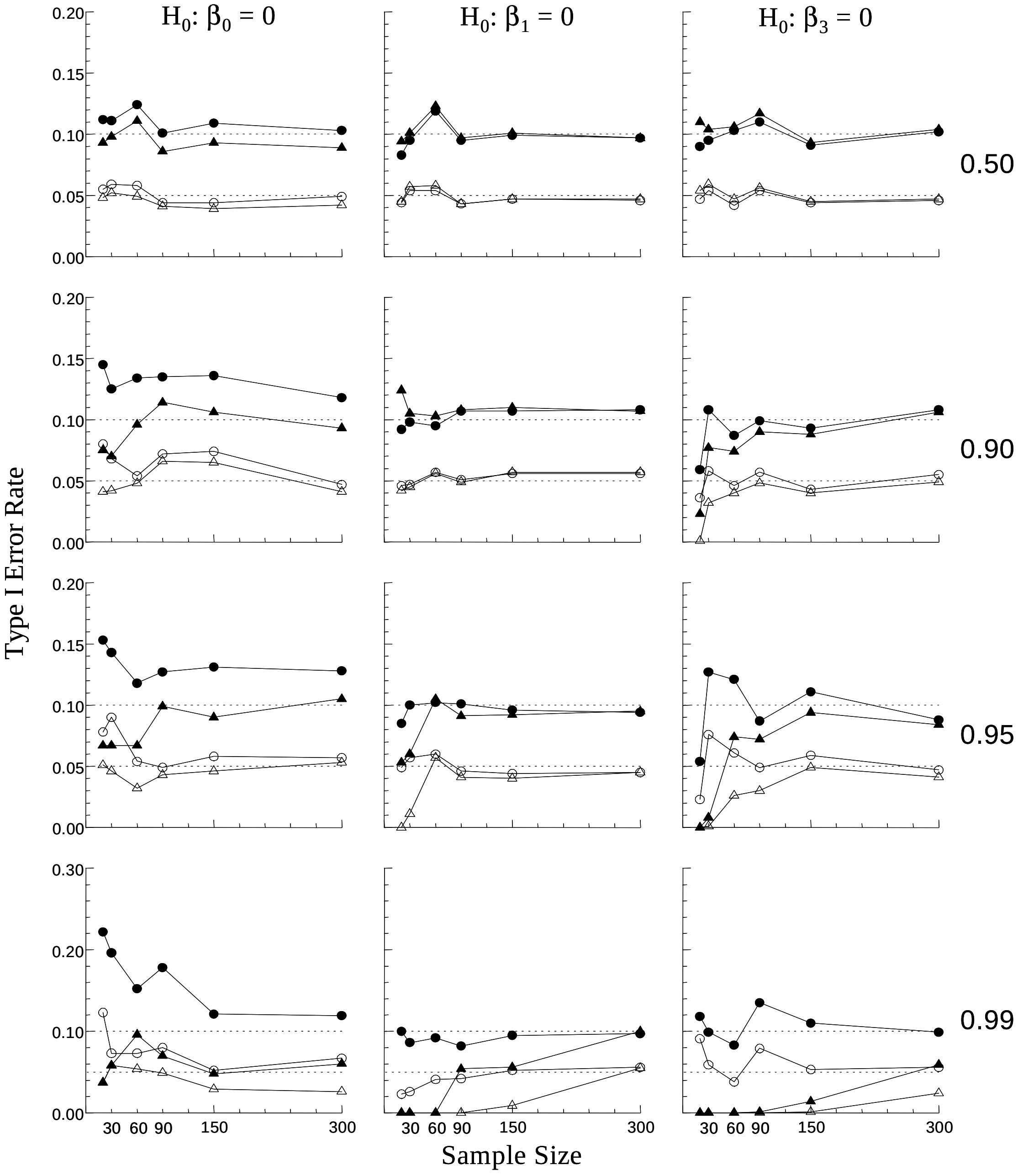
Turner, D. L., and D. C. Bowden. 1977. Simultaneous confidence bands for percentile lines in the general linear model. Journal of the American Statistical Association **72**:886-889.

Vardeman, S. B. 1992. What about the other intervals? The American Statistician **46**: 193-197.

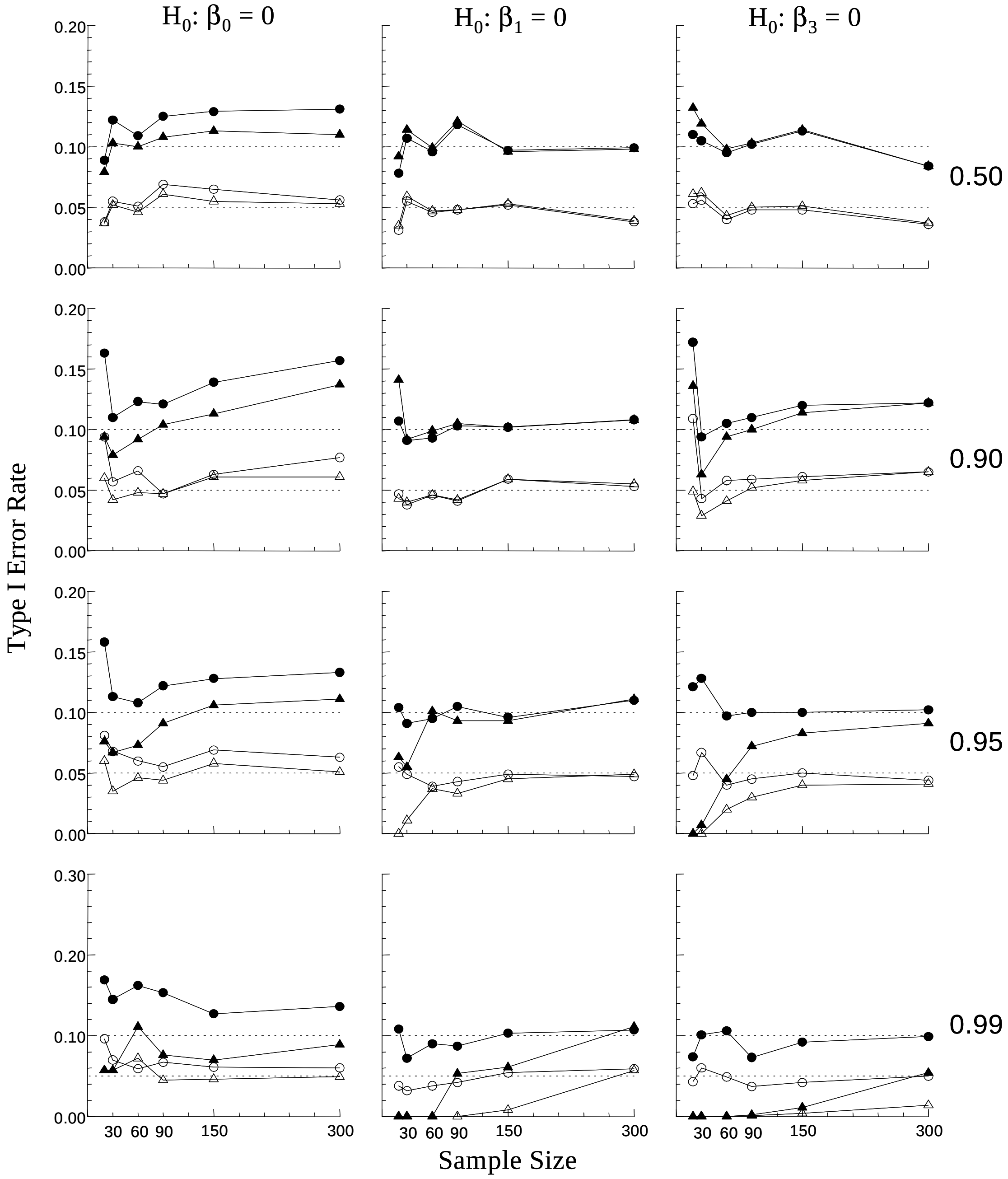
Zhou, K. G., and S. L. Portnoy. 1998. Statistical inference on heteroscedastic models based on regression quantiles. Nonparametric Statistics **9**:239-260.

Appendix 2

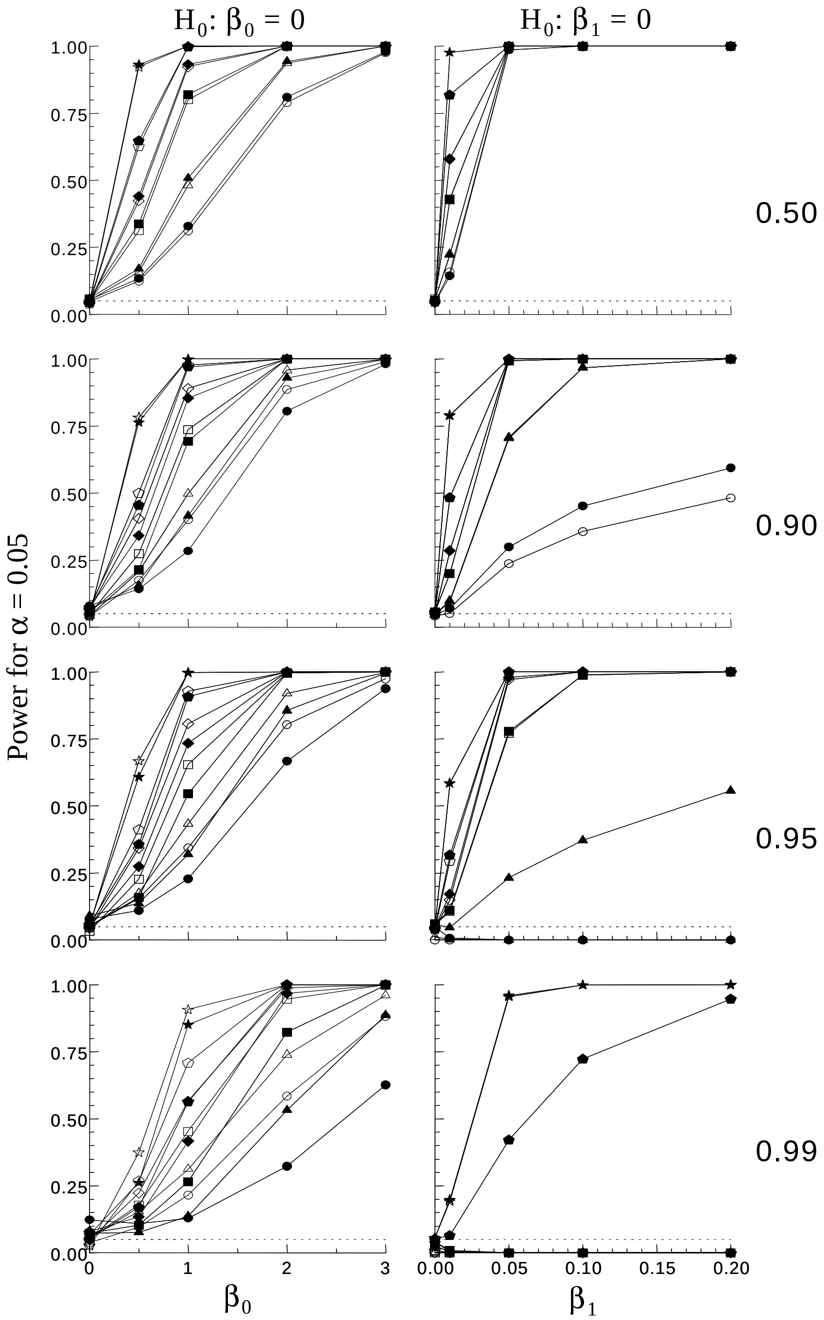
Simulation Results for Normal and Uniform Error Distributions



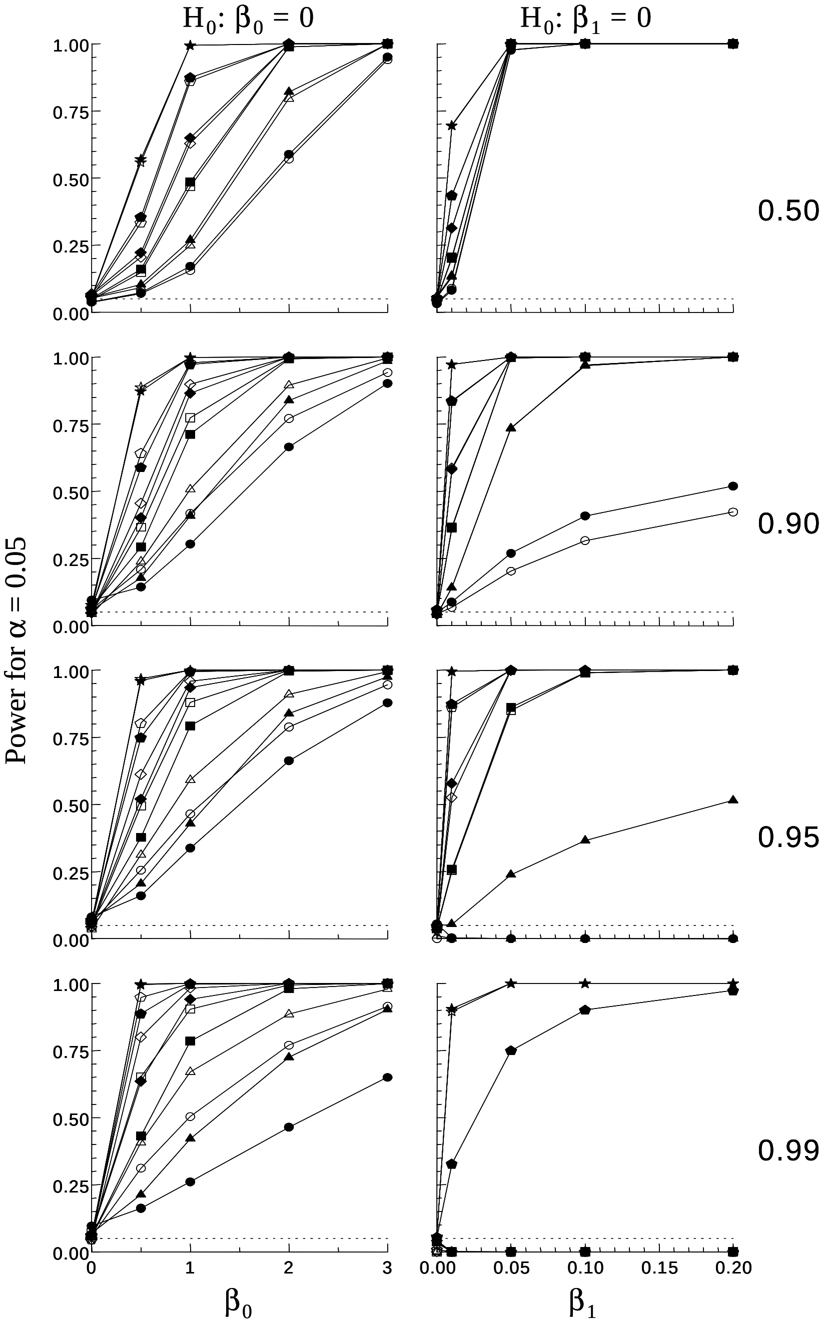
Appendix 2.1. Estimated type I error rates for $\alpha = 0.05$ (open) and 0.10 (solid); for the permutation *F* (circles) and Chi-square distributed *T* (triangles) rankscore tests; for homogeneous normal error distributions; for H_0: $\beta_0 = 0$ and H_0: $\beta_1 = 0$ in the model $y = \beta_0 + \beta_1 X_1 + \varepsilon$, and H_0: $\beta_3 = 0$ in the model $y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \varepsilon$; for 0.50, 0.90, 0.95, and 0.99 quantiles; and for n = 20, 30, 60, 90, 150, and 300. 1,000 random samples were used at each combination of H_0:, n, and quantile.



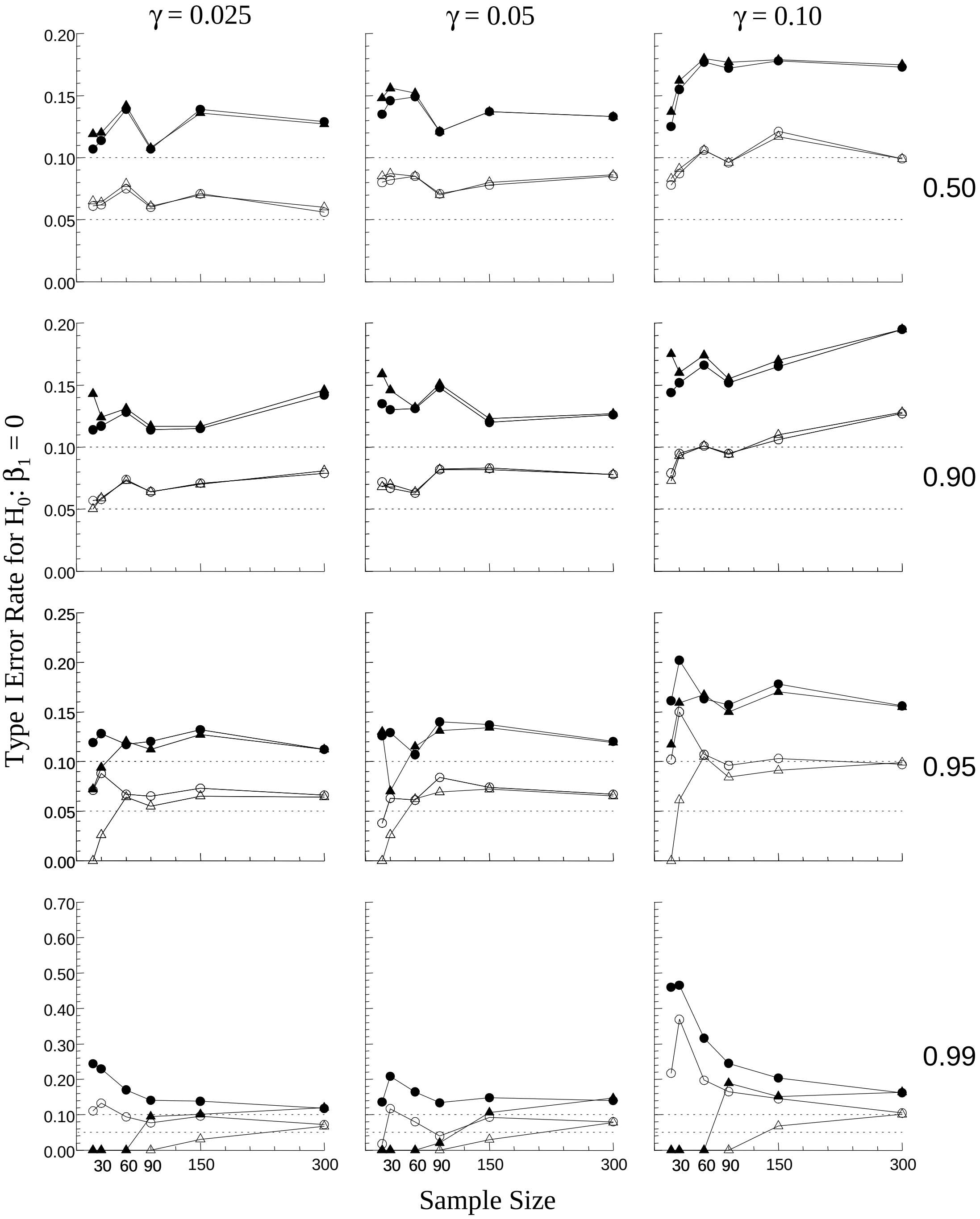
Appendix 2.2. Estimated type I error rates for $\alpha = 0.05$ (open) and 0.10 (solid); for the permutation F (circles) and Chi-square distributed T (triangles) rankscore tests; for homogeneous uniform error distributions; for H_0: $\beta_0 = 0$ and H_0: $\beta_1 = 0$ in the model $y = \beta_0 + \beta_1 X_1 + \varepsilon$, and H_0: $\beta_3 = 0$ in the model $y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \varepsilon$; for 0.50, 0.90, 0.95, and 0.99 quantiles; and for n = 20, 30, 60, 90, 150, and 300. 1,000 random samples were used at each combination of H_0:, n, and quantile.



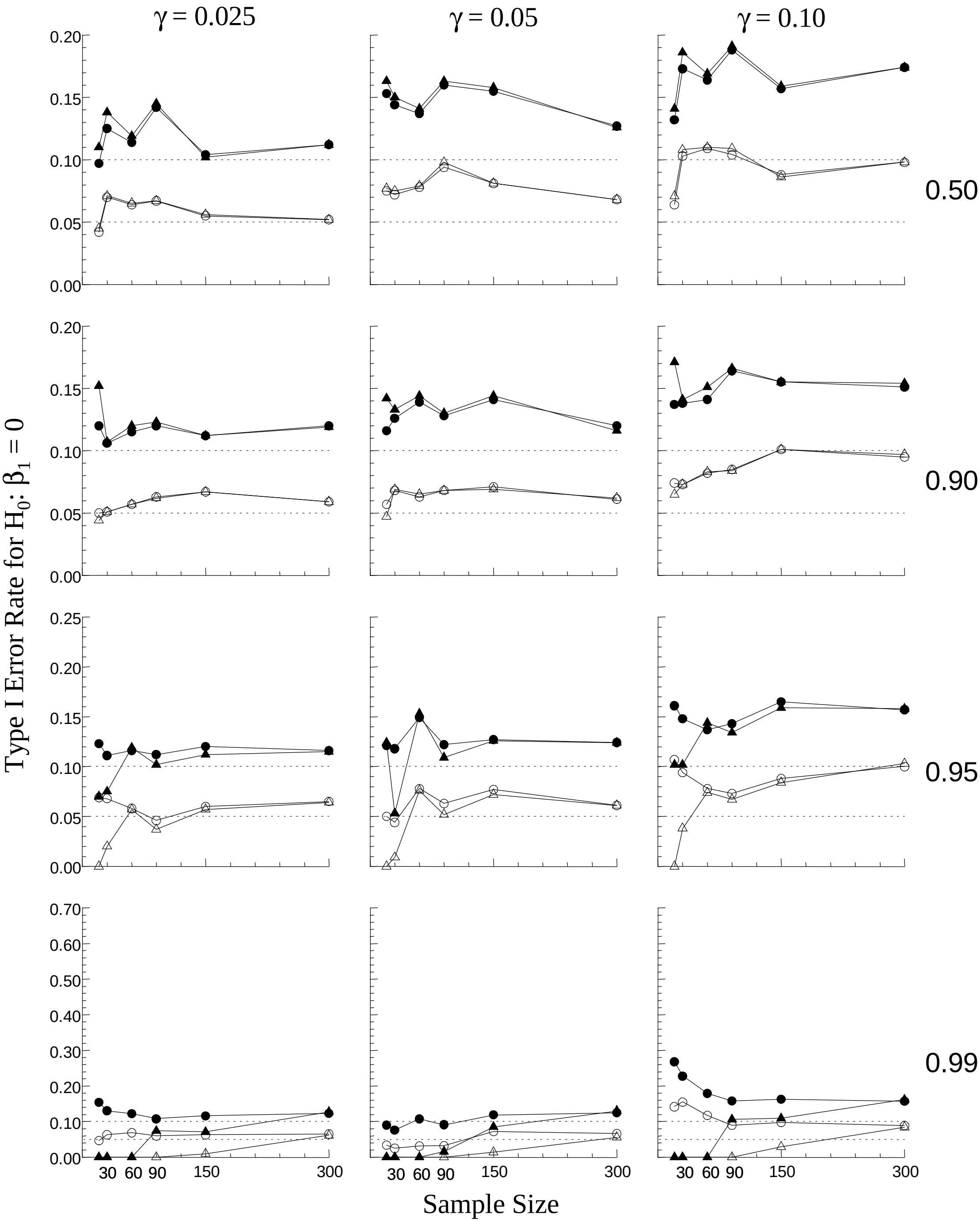
Appendix 2.3. Estimated power for $\alpha = 0.05$ for the permutation F (solid) and Chi-square distributed T (open) rankscore tests; for homogeneous normal error distributions; for H_0: $\beta_0 = 0$ and H_0: $\beta_1 = 0$ in the model $y = \beta_0 + \beta_1 X_1 + \varepsilon$; for β_0 = 0.0, 0.5, 1.0, 2.0, and 3.0 and for β_1 = 0.0, 0.01, 0.05, 0.10, and 0.20; for 0.50, 0.90, 0.95, and 0.99 quantiles; and for n = 20 (circle), 30 (triangle), 60 (square), 90 (diamond), 150 (pentagon), and 300 (star). Open symbols often are hidden behind solid symbols when equal. 1,000 random samples were used at each combination of effect size, n, and quantile.



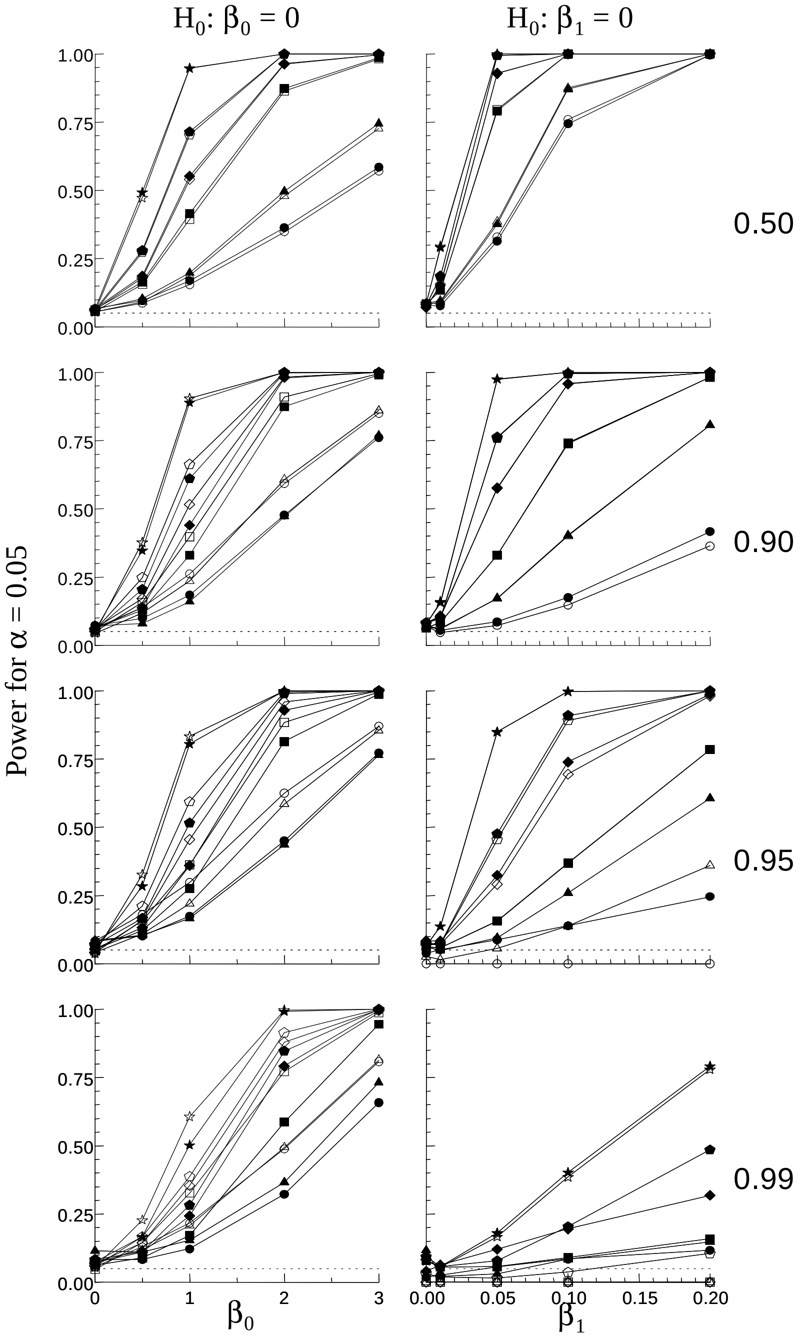
Appendix 2.4. Estimated power for $\alpha = 0.05$ for the permutation F (solid) and Chi-square distributed T (open) rankscore tests; for homogeneous uniform error distributions; for H_0: $\beta_0 = 0$ and H_0: $\beta_1 = 0$ in the model $y = \beta_0 + \beta_1 X_1 + \varepsilon$; for β_0 = 0.0, 0.5, 1.0, 2.0, and 3.0 and for β_1 = 0.0, 0.01, 0.05, 0.10, and 0.20; for 0.50, 0.90, 0.95, and 0.99 quantiles; and for n = 20 (circle), 30 (triangle), 60 (square), 90 (diamond), 150 (pentagon), and 300 (star). Open symbols often are hidden behind solid symbols when equal. 1,000 random samples were used at each combination of effect size, n, and quantile.



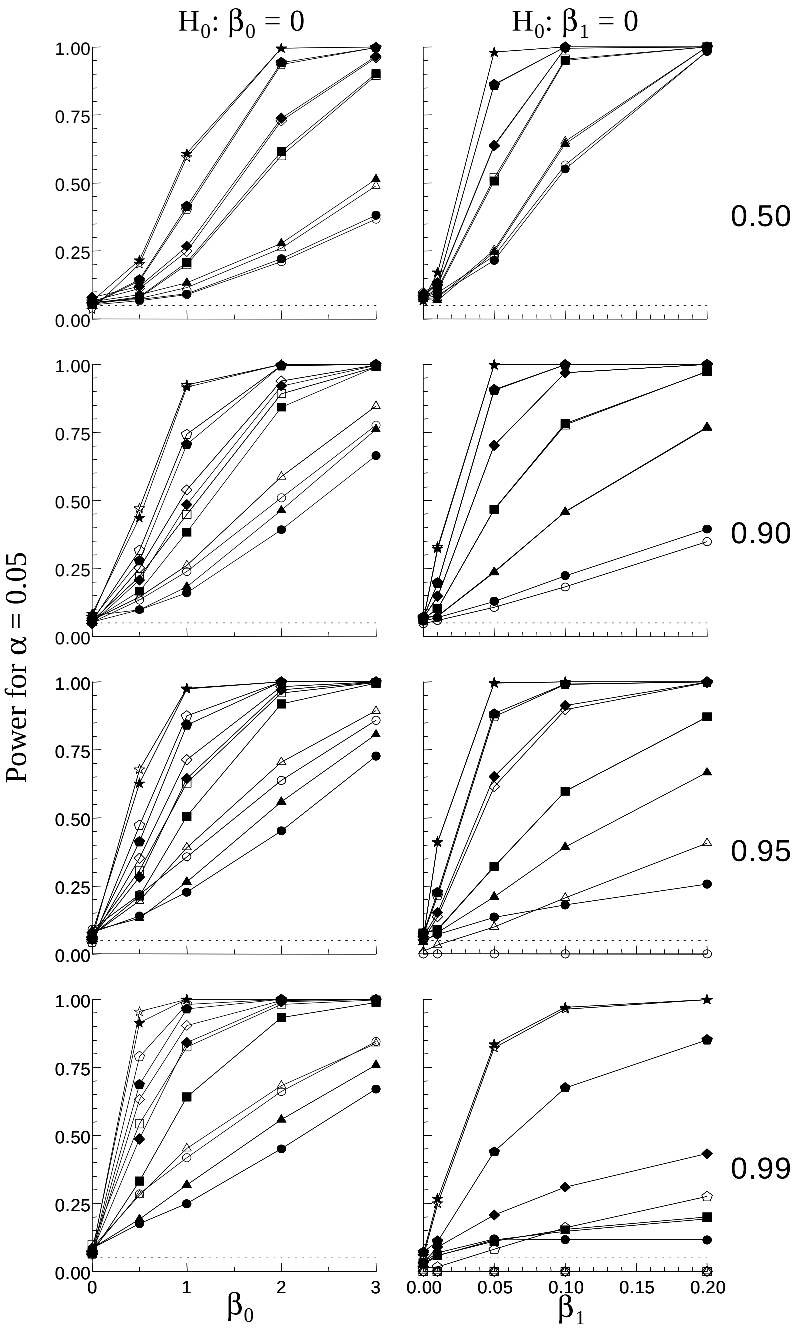
Appendix 2.5. Estimated type I error rates for $\alpha = 0.05$ (open) and 0.10 (solid); for the permutation *F* (circles) and Chi-square distributed *T* (triangles) rankscore tests for H_0: $\beta_1 = 0$; for heterogeneous normal error distributions with $\gamma = 0.025$, 0.05, and 0.10 in the model $y = \beta_0 + \beta_1 X_1 + (1 + \gamma X_1)\varepsilon$; for 0.50, 0.90, 0.95, and 0.99 quantiles; and for $n = 20$, 30, 60, 90, 150, and 300. 1,000 random samples were used at each combination of γ, n, and quantile.



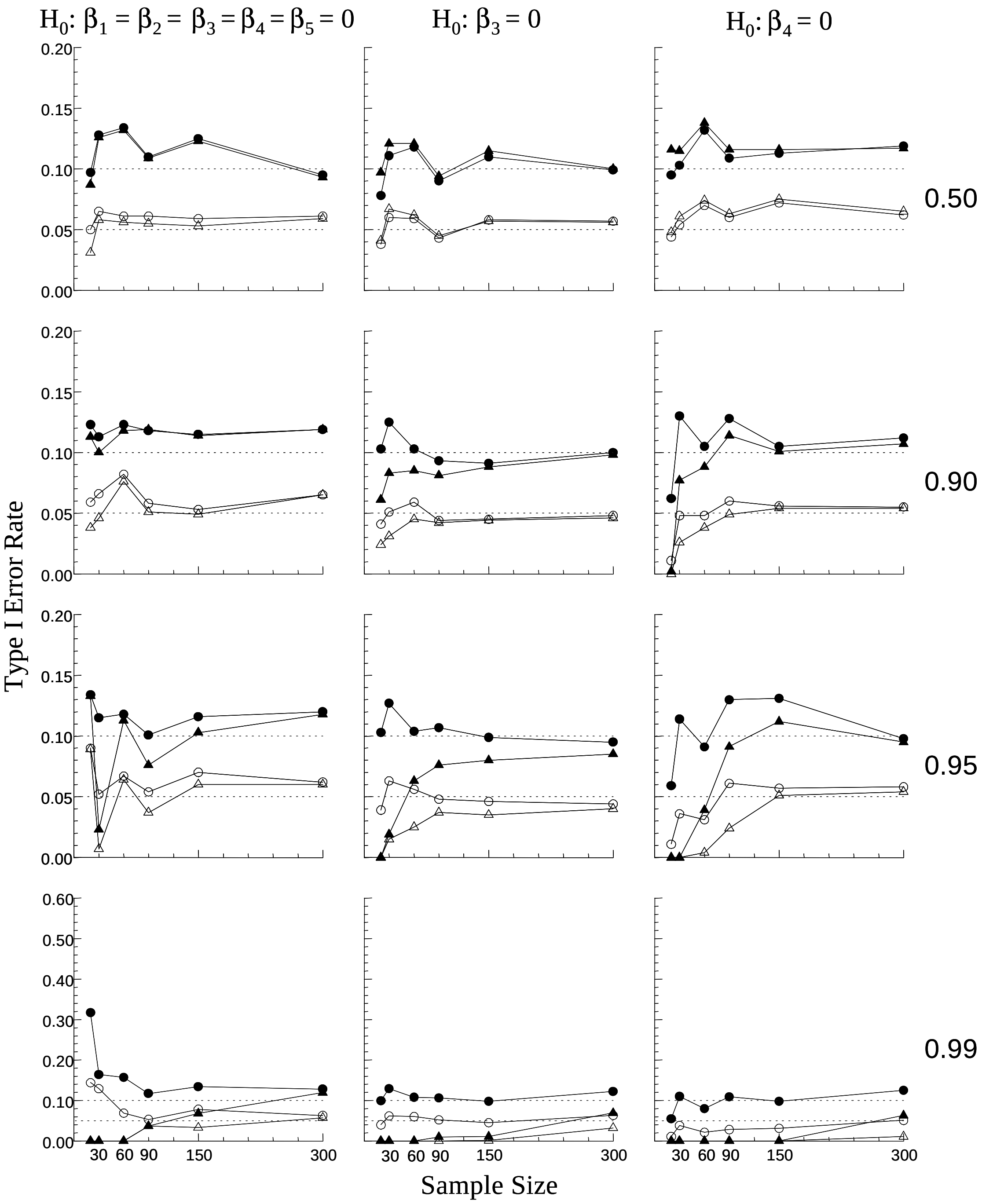
Appendix 2.6. Estimated type I error rates for $\alpha = 0.05$ (open) and 0.10 (solid); for the permutation *F* (circles) and Chi-square distributed *T* (triangles) rankscore tests for H_0: $\beta_1 = 0$; for heterogeneous uniform error distributions with $\gamma = 0.025$, 0.05, and 0.10 in the model $y = \beta_0 + \beta_1 X_1 + (1 + \gamma X_1)\varepsilon$; for 0.50, 0.90, 0.95, and 0.99 quantiles; and for $n = 20$, 30, 60, 90, 150, and 300. 1,000 random samples were used at each combination of γ, n, and quantile.



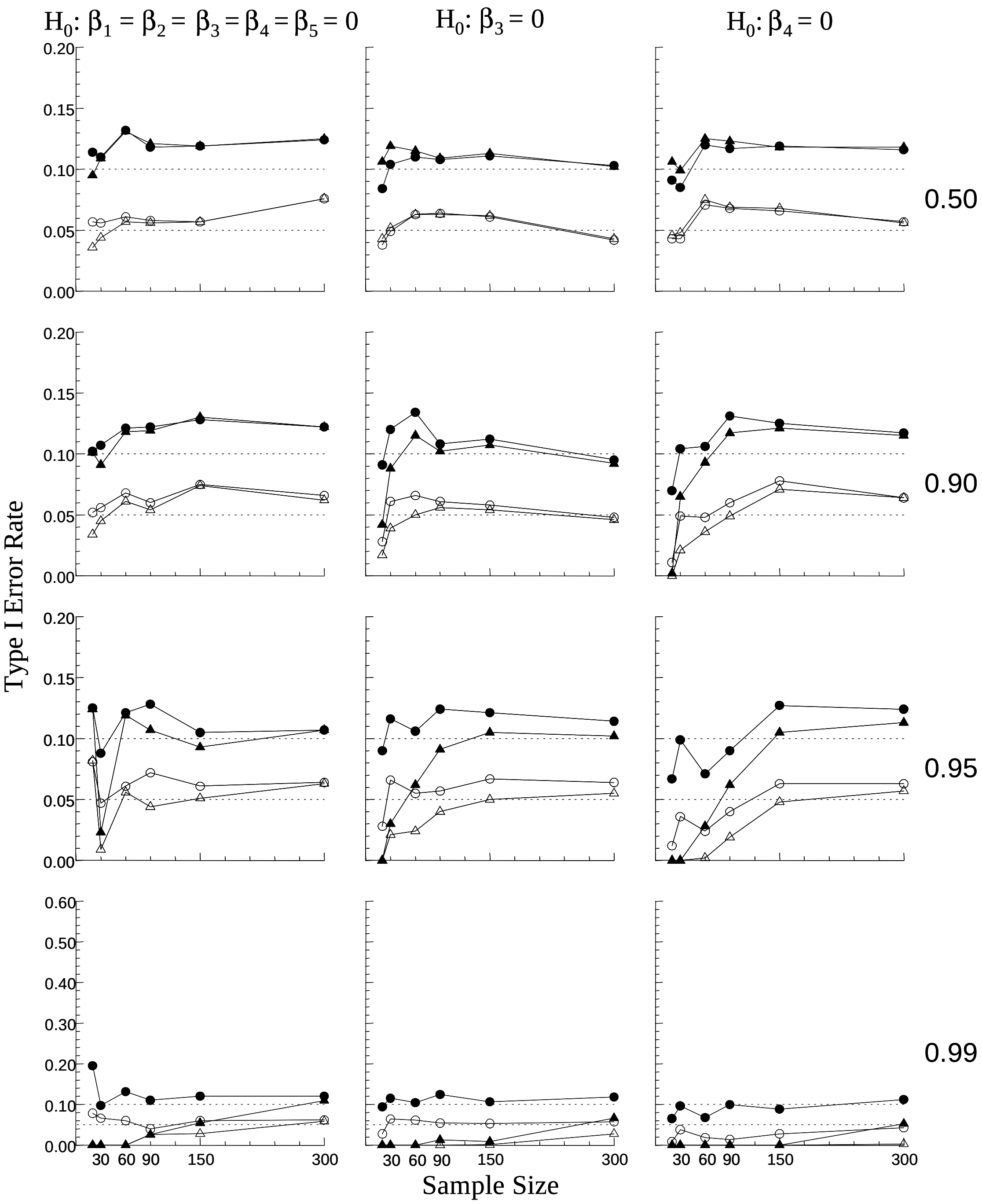
Appendix 2.7. Estimated power for $\alpha = 0.05$ for the permutation *F* (solid) and Chi-square distributed *T* (open) rankscore tests; for heterogeneous normal error distributions; for H_0: $\beta_0 = 0$ and H_0: $\beta_1 = 0$ in the model $y = \beta_0 + \beta_1 X_1 + (1 + 0.05X_1)\varepsilon$; for β_0 = 0.0, 0.5, 1.0, 2.0, and 3.0 and for β_1 = 0.0, 0.01, 0.05, 0.10, and 0.20; for 0.50, 0.90, 0.95, and 0.99 quantiles; and for *n* = 20 (circle), 30 (triangle), 60 (square), 90 (diamond), 150 (pentagon), and 300 (star). Open symbols often are hidden behind solid symbols when equal. 1,000 random samples were used at each combination of effect size, *n*, and quantile.



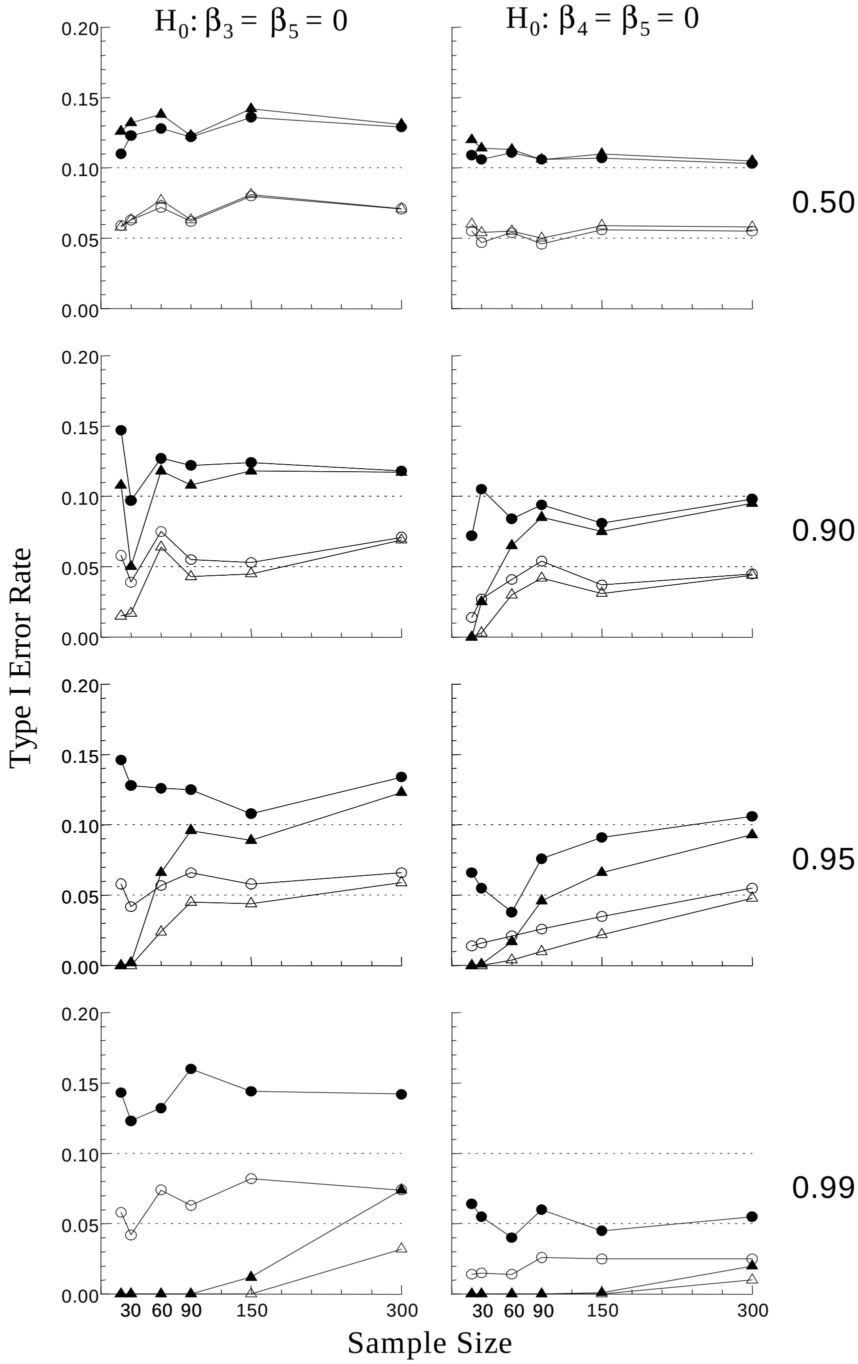
Appendix 2.8. Estimated power for $\alpha = 0.05$ for the permutation F (solid) and Chi-square distributed T (open) rankscore tests; for heterogeneous uniform error distributions; for H_0: $\beta_0 = 0$ and H_0: $\beta_1 = 0$ in the model $y = \beta_0 + \beta_1 X_1 + (1 + 0.05X_1)\varepsilon$; for β_0 = 0.0, 0.5, 1.0, 2.0, and 3.0 and for β_1 = 0.0, 0.01, 0.05, 0.10, and 0.20; for 0.50, 0.90, 0.95, and 0.99 quantiles; and for n = 20 (circle), 30 (triangle), 60 (square), 90 (diamond), 150 (pentagon), and 300 (star). Open symbols often are hidden behind solid symbols when equal. 1,000 random samples were used at each combination of effect size, n, and quantile.



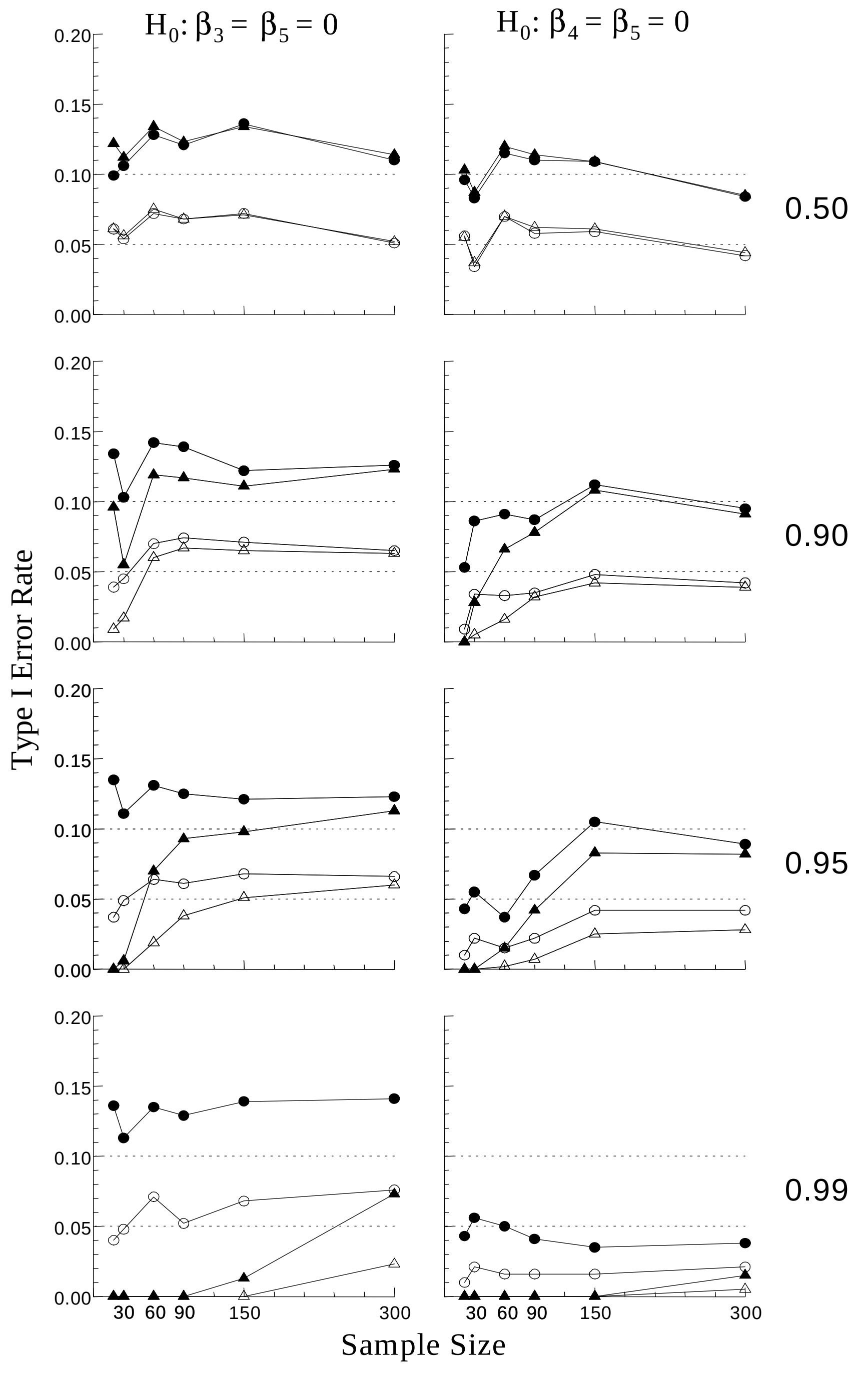
Appendix 2.9. Estimated type I error rates for $\alpha = 0.05$ (open) and 0.10 (solid); for the permutation F (circles) and Chi-square distributed T (triangles) rankscore tests for H_0: $\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0$, H_0: $\beta_3 = 0$, and H_0: $\beta_4 = 0$; for heterogeneous normal error distributions with $\gamma = 0.05$ in the model $y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + (1 + \gamma X_1)\varepsilon$; for 0.50, 0.90, 0.95, and 0.99 quantiles; and for n = 20, 30, 60, 90, 150, and 300. 1,000 random samples were used at each combination of H_0, n, and quantile.



Appendix 2.10. Estimated type I error rates for $\alpha = 0.05$ (open) and 0.10 (solid); for the permutation F (circles) and Chi-square distributed T (triangles) rankscore tests for H_0: $\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0$, H_0: $\beta_3 = 0$, and H_0: $\beta_4 = 0$; for heterogeneous uniform error distributions with $\gamma = 0.05$ in the model $y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + (1 + \gamma X_1)\varepsilon$; for 0.50, 0.90, 0.95, and 0.99 quantiles; and for n = 20, 30, 60, 90, 150, and 300. 1,000 random samples were used at each combination of H_0, n, and quantile.



Appendix 2.11. Estimated type I error rates for $\alpha = 0.05$ (open) and 0.10 (solid); for the permutation F (circles) and Chi-square distributed T (triangles) rankscore tests for H_0: $\beta_3 = \beta_5 = 0$ and H_0: $\beta_4 = \beta_5 = 0$ for heterogeneous normal error distributions with $\gamma = 0.05$ in the model $y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + (1 + \gamma X_1)\varepsilon$; for 0.50, 0.90, 0.95, and 0.99 quantiles; and for n = 20, 30, 60, 90, 150, and 300. 1,000 random samples were used at each combination of H_0, n, and quantile.



Appendix 2.12. Estimated type I error rates for α = 0.05 (open) and 0.10 (solid); for the permutation *F* (circles) and Chi-square distributed *T* (triangles) rankscore tests for H_0: $\beta_3 = \beta_5 = 0$ and H_0: $\beta_4 = \beta_5 = 0$ for heterogeneous uniform error distributions with $\gamma = 0.05$ in the model $y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + (1 + \gamma X_1)\varepsilon$; for 0.50, 0.90, 0.95, and 0.99 quantiles; and for n = 20, 30, 60, 90, 150, and 300. 1,000 random samples were used at each combination of H_0, n, and quantile.

Chapter 3

A Drop In Dispersion Permutation Test for Regression Quantile Estimates

Abstract: A drop in dispersion, *F*-ratio like permutation test for hypothesis testing and constructing confidence intervals for linear quantile regression estimates ($0 \leq \tau \leq 1$) was evaluated for conditions relevant to ecological investigations of animal responses to their physical environment. Conditions evaluated included models with 2 - 6 predictors, moderate collinearity among predictors, homogeneous and heterogeneous errors, small to moderate samples (n = 20 - 300), and central to upper quantiles (0.50 - 0.99). The drop in dispersion D test maintained Type I errors well for homogeneous error distributions and provided greater power than rankscore tests, which don't use the magnitude of the residuals in their construction. Type I errors for the D test were slightly liberal for weighted estimates of heterogeneous error distributions. The D test required larger n at more extreme quantiles than the rankscore tests to maintain reasonable Type I error rates and had more liberal Type I error rates when testing subhypotheses in multiple regression models. Confidence intervals on parameters were constructed based on test inversion for an example application relating trout densities to stream channel width:depth.

1. Introduction

Estimating the quantiles ($0 \leq \tau \leq 1$) of a response variable conditional on some set of covariates in a linear model has many applications in the biological and ecological sciences (Terrell et al. 1996, Scharf et al. 1998, Cade et al. 1999, Cade and Guo 2000, Haire et al. 2000, Eastwood et al. 2001, Dunham et al. 2002). Quantile regression models allow the entire conditional distribution of a response variable y to be related to some covariates X, providing a richer description of functional changes than is possible by focusing on just the mean (or other central statistics), yet requiring minimal distributional assumptions (Koenker and Bassett 1978, 1982, Koenker and Machado 1999). Quantile regression estimates are especially enlightening for relationships involving heterogeneous responses where by definition rates of change are not the same across all parts of the response distribution. Many ecological applications of quantile regression have focused on estimating some upper quantiles to characterize effects of limiting factors (Terrell et al. 1996, Scharf et al. 1998, Cade et al. 1999, Cade and Guo 2000, Haire et al. 2000, Eastwood et al. 2001, Huston 2002). Other applications (Allen et al. 2001, Dunham et al. 2002) have used estimates across the entire [0, 1] interval of quantiles as a flexible method of characterizing effects associated with heterogeneous distributions.

Inference methods with asymptotic validity for linear quantile regression models based on estimates of the covariance matrices (Koenker and Bassett 1978, 1982, Koenker and Machado 1999) require estimates of the reciprocal of the error density function at the quantile of interest, $f(F^{-1}(0))$. These methods often perform poorly at smaller sample sizes (Koenker 1987, Buchinsky 1991) and the asymptotic theory becomes suspect at more extreme (>0.7 and <0.3) quantiles (Chernozhukov and Umantsev 2001). Koenker (1994) introduced the idea of constructing confidence intervals by inverting a quantile rankscore test (Gutenbrunner et al. 1993), which does not require estimating the sparsity function, as an alternative inference procedure that performed well under linear heteroscedastic regression models and smaller sample sizes. Here I consider a drop in dispersion, *F*-ratio like test that is evaluated with permutation arguments based on modifications of the least absolute deviation regression test of Cade and Richards (1996). This test also avoids the sparsity estimation issue but unlike the quantile rankscore tests (Koenker 1994, Chapter 2) it uses the magnitude of the residuals in its construction, potentially providing greater power and shorter confidence intervals.

Here I evaluated performance of the drop in dispersion permutation test for central to extreme quantiles, a range of error structures, small to moderate sample sizes, and model forms likely to be encountered in ecological applications where the objective is to estimate some organism's response to its environment. Weighted forms of the test based on weighted quantile regression estimates were evaluated for heterogeneous error distributions. The drop in dispersion permutation test was applied to a quantile regression analysis of Lahontan cutthroat trout (*Oncorhynchus clarki henshawi*) response to variations in their stream habitat, expanding on the previous analyses of Dunham et al. (2002).

2. Quantile Regression Model

The τ^{th} regression quantile ($0 \leq \tau \leq 1$) for the heteroscedastic linear location-scale model $\mathbf{y} = \mathbf{X\beta} + \mathbf{\Gamma\varepsilon}$ is defined as $Q_y(\tau|\mathbf{X}) = \mathbf{X\beta}(\tau)$ and $\boldsymbol{\beta}(\tau) = \boldsymbol{\beta} + F_\varepsilon^{-1}(\tau)\boldsymbol{\gamma}$; where $\mathbf{y}$ is an $n \times 1$ vector of dependent responses, $\boldsymbol{\beta}$ is a $p \times 1$ vector of unknown regression parameters, $\mathbf{X}$ is an $n \times p$ matrix of predictors (first column consists of 1's for an intercept term), $\boldsymbol{\gamma}$ is a $p \times 1$ vector of unknown scale parameters, $\mathbf{\Gamma}$ is a diagonal $n \times n$ matrix where the n diagonal elements are the n corresponding ordered elements of the $n \times 1$ vector $\mathbf{X\gamma}$ (diag($\mathbf{X\gamma}$)), $\boldsymbol{\varepsilon}$ is an $n \times 1$ vector of random errors that are independent and identically distributed (iid), and F_ε^{-1} is the inverse of the cumulative distribution of the errors (Koenker and Bassett 1982, Buchinsky 1991, Gutenbrunner and Jurečková 1992, Koenker and Machado 1999). Homoscedastic regression models are a special case of the linear-location scale model when $\boldsymbol{\gamma} = (1,0,...,0)'$ and $Q_y(\tau|\mathbf{X}) = \mathbf{X\beta}(\tau)$, $\boldsymbol{\beta}(\tau) = \boldsymbol{\beta} + (F_\varepsilon^{-1}(\tau),0,...,0)'$, where all parameters other than the intercept (β_0) in $\boldsymbol{\beta}(\tau)$ are the same for all τ. More general forms of heteroscedastic errors can be accommodated with regression quantiles (Koenker 1997, Koenker and Machado 1999) but were not considered here.

The restriction imposed on F_ε to estimate regression quantiles is that a τ^{th} quantile of $\mathbf{y} - \mathbf{X\beta}(\tau)$ conditional on $\mathbf{X}$ equals 0, $F_\varepsilon^{-1}(\tau|\mathbf{X}) = 0$. Estimates, $\mathbf{b}(\tau)$, of $\boldsymbol{\beta}(\tau)$ are solutions to the following minimization problem:

$$\min\Big[\sum_{i=1}^{n} \rho_\tau\Big(y_i - \sum_{j=0}^{p} b_j x_{ij}\Big)\Big]$$
$$\textit{where } \rho_\tau(e) = e(\tau - I(e<0)),$$
$$\textit{and } I(\cdot) \textit{ is the indicator function.} \tag{1}$$

The estimating equations in (1) yield primal solutions in a modification of the Barrodale and Roberts (1974) simplex linear program for any specified value of τ (Koenker and d'Orey 1987). With little additional computation the entire regression quantile process for all distinct values of τ can be estimated (Koenker and d'Orey 1987, 1994).

Consistent estimates with reduced sampling variation for heteroscedastic linear models can be obtained by implementing weighted versions of the regression quantile estimators, where weights are based on the sparsity function at a given quantile and covariate value (Koenker and Portnoy 1996, Koenker and Machado 1999). In the linear location-scale model this simplified to using an $n \times n$ weights matrix, $\mathbf{W} = \mathbf{\Gamma}^{-1}$, where the $p \times 1$ vector of scale parameters $\boldsymbol{\gamma}$ would usually have to be estimated in applications (Gutenbrunner and Jurečková 1992, Koenker and Zhao 1994, Koenker and Machado 1999). The weighted regression quantile estimates then are given by

$$\begin{gathered} \min\Big[\sum_{i=1}^{n} \rho_\tau\Big(y_i - \sum_{j=0}^{p} b_j x_{ij}\Big) w_i\Big] \\ \textit{where } \rho_\tau(e) = e(\tau - I(e<0)), \\ w_i \textit{ is a weight}, \\ \textit{and } I(\cdot) \textit{ is the indicator function}. \end{gathered} \tag{2}$$

which is easily implemented by multiplying $\mathbf{y}$ and $\mathbf{X}$ by $\mathbf{W}$ and then using the unweighted estimator (1).

3. Test Statistics

The drop in dispersion *D* test was based on a modification of the drop in dispersion permutation test developed for least absolute deviation (LAD) regression (Cade and Richards 1996). The reduced parameter model, $\mathbf{y} - \mathbf{X}_2\boldsymbol{\xi}(\tau) = \mathbf{X}_1\boldsymbol{\beta}_1(\tau) + \mathbf{\Gamma}\boldsymbol{\varepsilon}$, is constructed by partitioning $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2)$, where $\mathbf{X}_1$ is $n \times (p - q)$ and $\mathbf{X}_2$ is $n \times q$; and by partitioning $\boldsymbol{\beta} = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2)$, where $\boldsymbol{\beta}_1(\tau)$ is a $(p - q) \times 1$ vector of unknown nuisance parameters under the null and $\boldsymbol{\beta}_2(\tau)$ is a $q \times 1$ vector of parameters specified by the null hypothesis H_0: $\boldsymbol{\beta}_2(\tau) = \xi(\tau)$ (frequently $\xi(\tau) = \mathbf{0}$) for the full parameter model $\mathbf{y} = \mathbf{X}_1\boldsymbol{\beta}_1(\tau) + \mathbf{X}_2\boldsymbol{\beta}_2(\tau) + \boldsymbol{\Gamma}\boldsymbol{\varepsilon}$; and $\mathbf{y}$, $\boldsymbol{\Gamma}$, and $\boldsymbol{\varepsilon}$ are as above. The sum of weighted absolute deviations minimized in (1) for the weighted version of the full parameter model, $\mathbf{Wy} = \mathbf{WX}_1\boldsymbol{\beta}_1(\tau) + \mathbf{WX}_2\boldsymbol{\beta}_2(\tau) + \mathbf{W\Gamma e}$, where $\mathbf{W}$ is a weights matrix as in (2), are denoted SAFw(τ) and for the reduced parameter model, $\mathbf{Wy} - \mathbf{WX}_2\xi(\tau) = \mathbf{WX}_1\boldsymbol{\beta}_1(\tau) + \mathbf{W\Gamma e}$, corresponding to the restrictions under the null hypothesis H_0: $\boldsymbol{\beta}_2(\tau) = \xi(\tau)$ are denoted SARw(τ). The test statistic

$$D_o = (\mathrm{SARw}(\tau) - \mathrm{SAFw}(\tau))/\mathrm{SAFw}(\tau), \tag{3}$$

was evaluated by permuting the weighted residuals under the null model to the weighted full model matrix $\mathbf{WX}$, similar to the Cade and Richards (1996) procedure. By taking a large random sample m of the $n!$ possible permutations, probability under the null hypothesis that $D \geq D_o$ was approximated by (the number of $D \geq D_o$ + 1)/(m + 1). When the error distributions are assumed homogeneous so that $\mathbf{W} = \mathbf{I}$, where $\mathbf{I}$ is the $n \times n$ identity matrix, and $\tau = 0.5$, this test statistic is identical to the statistic of Cade and Richards (1996) for LAD regression. The weights, $\mathbf{W}$, serve to eliminate the effects of heterogenous errors so that permuting weighted residuals provide an approximation of the sampling distribution of D.

Permuting residuals ($\mathbf{e} = \mathbf{y} - \mathbf{X}_1\mathbf{b}_1$) under the null reduced parameter model does not in general yield exact permutation probabilities except when the null parameter is just an intercept (β_0), but this permutation approach due to Freedman and Lane (1983) was found to have perfect correlation asymptotically with the exact test (only possible when $\boldsymbol{\beta}_1$ is known) (Anderson and Robinson 2001) and has performed well in simulation studies for least squares (Kennedy and Cade 1996, Anderson and Legendre 1999, Legendre 2000) and least absolute deviation regression (Cade and Richards 1996). There is some correlation ($-(n - 1)^{-1}$) among the residuals and they don't have constant variance ($E[\mathbf{e}\,\mathbf{e}'] = \sigma^2(\mathbf{I} - \mathbf{X_1}(\mathbf{X_1}'\mathbf{X_1})^{-1}\mathbf{X_1}')$), implying that they are not exactly exchangeable. Dependency among the residuals decreases with increasing sample size providing some asymptotic justification for treating them as exchangeable random variables (Randles 1984). Commenges (In Press) discusses transformations to preserve exchangeability of the first two moments of the residuals but those were not pursued here.

4. Simulation Experiment

I first conducted a set of Monte Carlo simulations with homogeneous errors to establish performance for models with simple error structure. Normal ($\mu = 0$, $\sigma = 1$), uniform (min = -2, max = 2), and lognormal (median = 0, $\sigma = 0.75$) error distributions were used to provide responses with symmetric, unimodal variation with greatest density at the center, symmetric variation with constant density, and asymmetric variation with density in a long upper tail. Error distributions were centered on their 0.50, 0.75, 0.90, 0.95, or 0.99 quantiles so that $F_\varepsilon^{-1}(\tau|\mathbf{X}) = 0$, providing a range of central to extreme regression quantiles. Note that similar simulation results for quantiles in the lower tail

(0.25, 0.10, 0.05, and 0.01) would be obtained for the symmetric error distributions.

Simple 2 parameter and 6 parameter multiple regression models were simulated for n = 20, 30, 60, 90, 150, and 300. Independent variables were structured to have a range of values and correlation structure similar to what might be expected in measures of forest habitat structure for avian species. Independent variables were structured so that X_0 was a column of 1's for the intercept; X_1 was uniformly distributed (0, 100); X_2 was negatively correlated (r = -0.89) with X_1 specified by the function $X_2 = 4{,}000 - 20X_1 + N(\mu = 0, \sigma = 300)$; X_3 was positively correlated (r = 0.94) with X_1 specified by the function $X_3 = 10 + 0.4X_1 + N(\mu = 0, \sigma = 16)$; X_4 was a 0,1 indicator variable randomly assigning half the sample to each of 2 groups; and X_5 was the multiplicative interaction of X_3 and X_4. Thus, X_1 ranged from 0 - 100 similar to measures of percent tree canopy cover, X_2 had most values in the range 0 - 5,000 and was inversely related to tree cover similar to density (stems/ha) of a shade intolerant shrub, and X_3 had most values in the range 0 - 60 similar to tree height (m) and was positively related to tree cover. Variables X_2 and X_3 were negatively correlated (r = -0.85) with each other through their indirect functional relation with X_1. The indicator variable (X_4) and its interaction with X_3 (X_5) allowed the effect of X_3 for the regression quantile function to differ in slopes, intercepts, or both terms for the 2 groups.

Each combination of conditions (quantile, error distribution, sample size, and model structure) was sampled 1,000 times, and the test statistic D_o was computed for each sample. Probabilities for the D test were evaluated with separate $m + 1 = 10{,}000$ random samples of the permutation distribution. Cumulative distribution function (cdf) plots of the Type I error probabilities under the null hypothesis were graphed and compared with the expected uniform cdf. However, point estimates for $\alpha = 0.05$ and 0.10, corresponding to coverage for 95% and 90% confidence intervals, were graphed across the combination of model conditions because the number of graphs required to display the cdf plots was excessive. The 99% binomial confidence interval for 1,000 simulations for $\alpha = 0.10$ is 0.076 - 0.124 and for $\alpha = 0.05$ is 0.032 - 0.068, which can be used as a guide to judge how much the estimated Type I error rates exceeded variation expected from the sampling simulations. Power under the alternative hypotheses was graphed only for $\alpha = 0.05$ across all combinations of conditions, although cdf plots were initially examined.

All data for the simulation studies were generated with functions in S-Plus 2000 (Mathsoft, Inc., Seattle, WA). Regression quantile estimates and test statistics were computed by a static memory compilation of Fortran 95 routines implemented in the Blossom software available from the U. S. Geological Survey (*www.mesc.usgs.gov/products/software/ blossom.shtml*).

4.1 Homogeneous Error Structure - Simple Regression

The simple 2 parameter regression model, $y = \beta_0 + \beta_1 X_1 + \varepsilon$ was evaluated for H_0: $\beta_1 = 0$ with β_0 fixed at 6.0 and β_1 = 0.0, 0.01, 0.05, 0.10, and 0.20. Type I error rates were well maintained at all sample sizes, error distributions, and quantiles, consistent with exact exchangeability for this hypothesis (Fig 3.1). Type I errors for the 0.75 quantile were nearly identical to those for the 0.50 quantile and, therefore, were

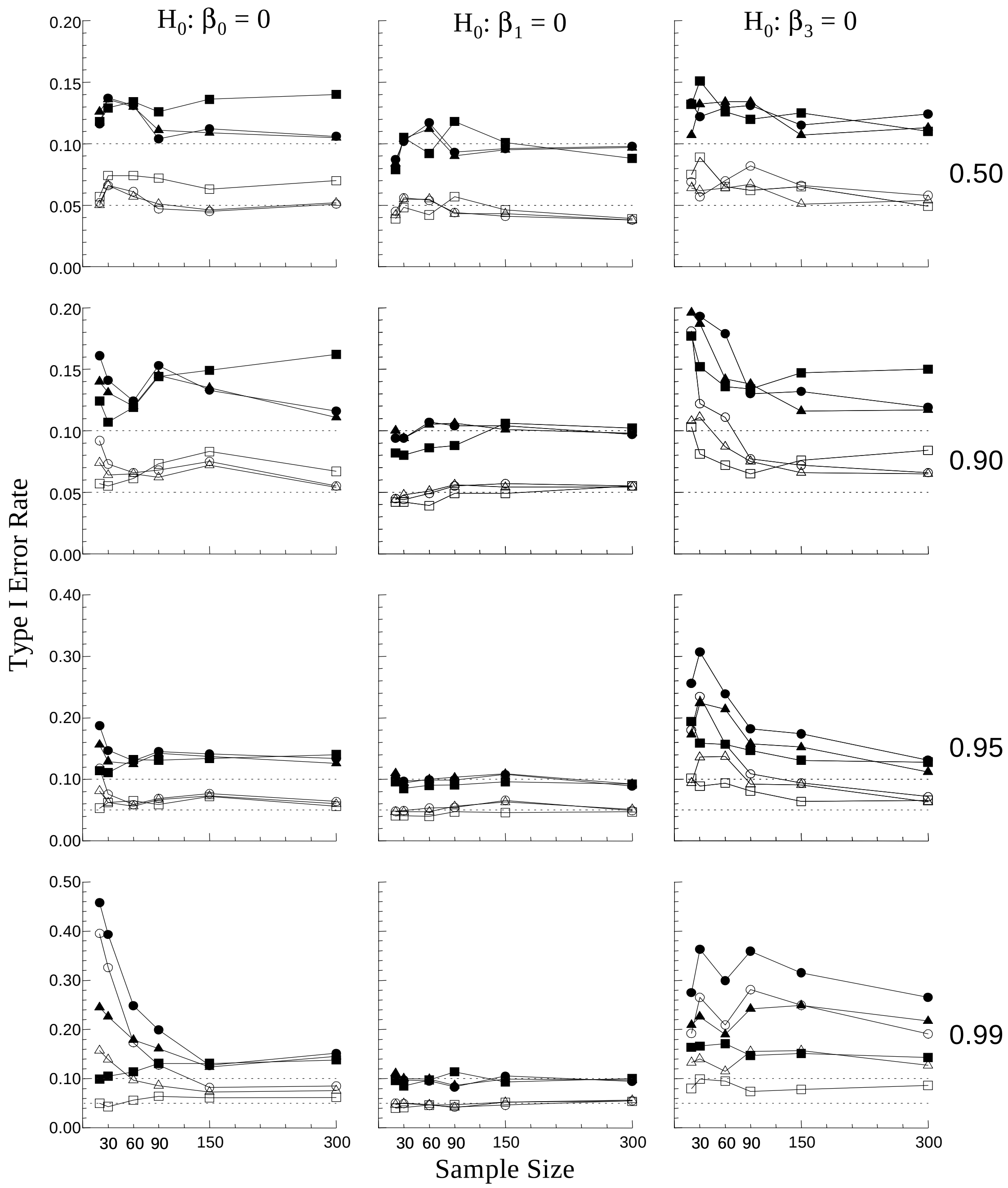


Figure 3.1. Estimated type I error rates for $\alpha = 0.05$ (open) and 0.10 (solid); for the permutation *D* test for homogeneous lognormal (circles), normal (triangles), and uniform (squares) error distributions; for H_0: $\beta_0 = 0$ and H_0: $\beta_1 = 0$ in the model $y = \beta_0 + \beta_1 X_1 + \varepsilon$, and H_0: $\beta_3 = 0$ in the model $y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \varepsilon$; for 0.50, 0.90, 0.95, and 0.99 quantiles; and for n = 20, 30, 60, 90, 150, and 300. 1,000 random samples were used at each combination of H_0:, n, and quantile.

not graphed for this or subsequent simulations. This regression model also was evaluated for H_0: $\beta_0 = 0$ with β_1 fixed at 0.10 and β_0 = 0.0, 0.5, 1.0, 2.0, and 3.0. Type I error rates for the intercept under then null hypothesis (β_0 = 0.0) were slightly liberal for all quantiles, becoming extremely liberal for $n < 150$ for 0.99 quantile (Fig. 3.1). A comparison of the cdf's for these two hypotheses for the lognormal error distribution and $n = 90$ provides another view of the degree to which H_0: $\beta_0 = 0$ deviates from the exactness of H_0: $\beta_1 = 0$ (Fig 3.2).

Power to detect nonzero slopes (β_1 = 0.01, 0.05, 0.10, 0.20) was progressively lower moving from the 0.50 to 0.99 quantile and this reduction was greatest for the lognormal error distribution (Fig. 3.3), less for the normal and least for the uniform error distributions (Appendices 3.1 and 3.2). For this and subsequent power simulations, the lognormal error distributions had lowest power and are given in the Figures; power for normal and uniform error distribution are in Appendix 3. Power for the *D* test was greater (relative power = 1.00 - 1.41) for 0.50 and 0.75 quantiles to much greater (relative power =1.00 - 4.91) for 0.90 - 0.99 quantiles than for the rankscore tests (Chapter 2). Power to detect nonzero intercepts (β_0 = 0.5, 1.0, 2.0, and 3.0) followed a similar reduction with increasing quantile for the lognormal error distribution, with no effective power for the 0.99 quantile and $n < 150$ (Fig. 3.3). The normal and uniform error distributions had less reduction in power across quantiles for this hypothesis and had effective power for all samples sizes for the 0.99 quantile (Appendices 3.1 and 3.2). The rankscore tests (Chapter 2) had power comparable to the *D* test for this hypothesis.

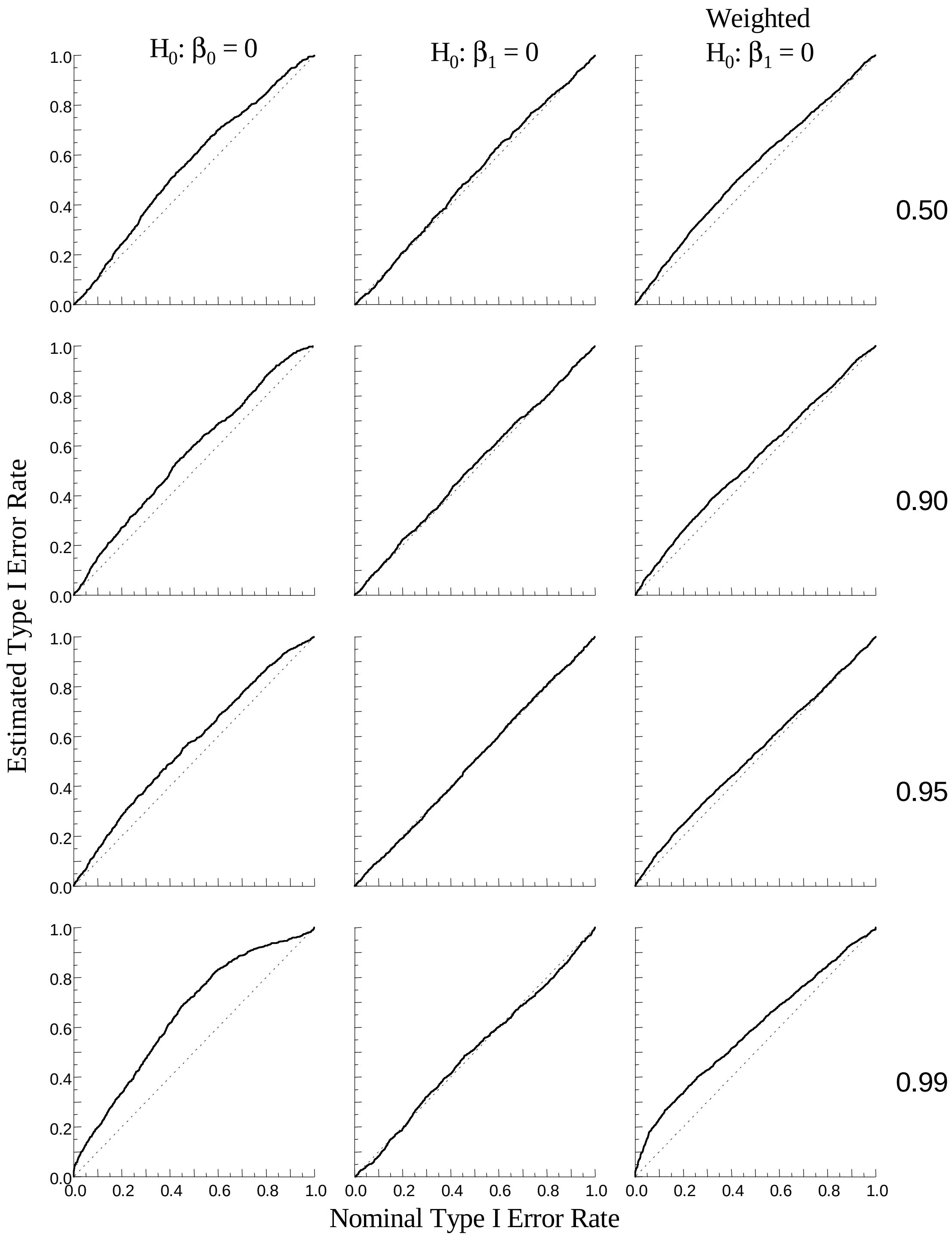


Figure 3.2. Cumulative distributions of 1,000 estimated errors for permutation D tests of H_0: $\beta_0 = 0$ and H_0: $\beta_1 = 0$ for the model $y = \beta_0 + \beta_1 X_1 + e$, and H_0: $\beta_1 = 0$ for the weighted model $wy = w\beta_0 + w\beta_1 X_1 + w(1 + \gamma X_1)\varepsilon$ with $\gamma = 0.05$ and $w = (1 + \gamma X_1)^{-1}$; for 0.50, 0.90, 0.95, and 0.99 quantiles for the lognormal error distribution and $n = 90$.

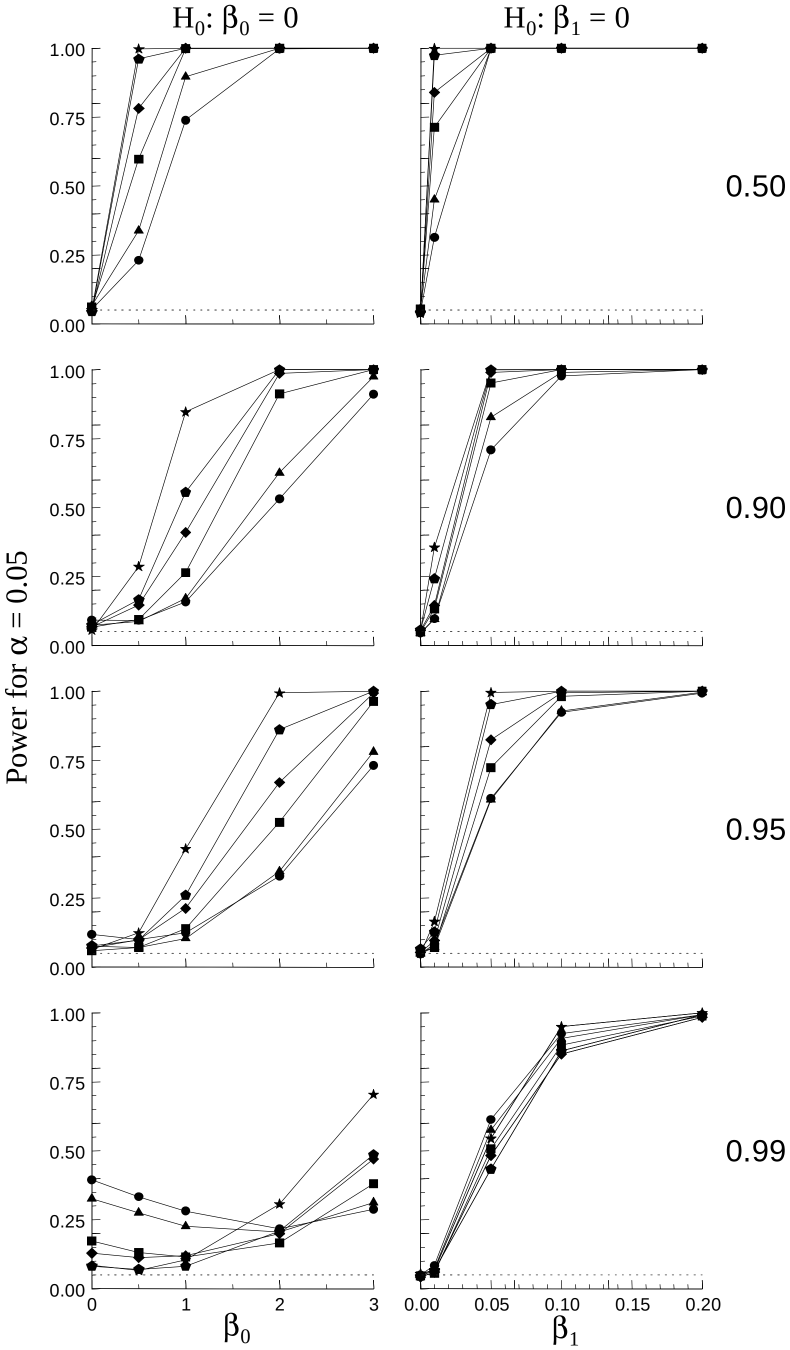


Figure 3.3. Estimated power for $\alpha = 0.05$ for the permutation D tests for homogeneous lognormal error distributions for H_0: $\beta_0 = 0$ and H_0: $\beta_1 = 0$ in the model $y = \beta_0 + \beta_1 X_1 + \varepsilon$; for $\beta_0 = 0.0$, 0.5, 1.0, 2.0, and 3.0 and for $\beta_1 = 0.0$, 0.01, 0.05, 0.10, and 0.20; for 0.50, 0.90, 0.95, and 0.99 quantiles; and for $n = 20$ (circle), 30 (triangle), 60 (square), 90 (diamond), 150 (pentagon), and 300 (star). 1,000 random samples were used at each combination of effect size, n, and quantile.

4.2 Homogeneous Error Structure - Multiple Regression

The 6 parameter model, $y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \varepsilon$, was evaluated for H_0: $\beta_3 = 0$ with $\beta_0 = 36.0$, $\beta_1 = 0.10$, $\beta_2 = -0.005$, $\beta_4 = 2.0$, and $\beta_3 = \beta_5 = 0.0$. Type I error rates were always slightly liberal, becoming more liberal with increasing quantile and decreasing sample size to the point that error rates were totally unreliable for the 0.99 quantile (Fig. 3.1). Type I error rates for the uniform error distribution did not degrade as much as those for the lognormal and normal error distributions. The 6 parameter model also was evaluated for H_0: $\beta_4 = 0$ with $\beta_0 = 36.0$, $\beta_1 = 0.10$, $\beta_2 = -0.005$, $\beta_3 = 0.05$, and $\beta_4 = \beta_5 = 0.0$. A similar pattern of liberal Type I error rates was found. Power was not investigated for multiple regression models with homogeneous errors.

4.3 Heterogeneous Error Structure - Simple Regression

The 2 parameter weighted regression model with heterogeneous errors, $wy = w\beta_0 + w\beta_1 X_1 + w(1 + \gamma X_1)\varepsilon$, was evaluated with $\gamma = 0.05$ using the known weights $w = (1 + 0.05X_1)^{-1}$ for H_0: $\beta_1 = 0$ with $\beta_0 = 6.0$ and $\beta_1 = 0.0$. Type I error rates were slightly liberal for 0.50 - 0.90 quantiles, becoming increasingly liberal from 0.95 to 0.99 quantiles with decreasing sample size (Fig. 3.4). Type I error rates were not as liberal for uniform compared to lognormal and normal error distributions at higher quantiles and smaller sample sizes. Examining the cdf's for this hypothesis for the lognormal error distribution and $n = 90$ provides another view of the degree to which H_0: $\beta_1 = 0$ for the weighted estimate in the heterogeneous error distribution model deviates from the exactness of H_0: $\beta_1 = 0$ in the homogeneous error distribution model (Fig 3.2). The

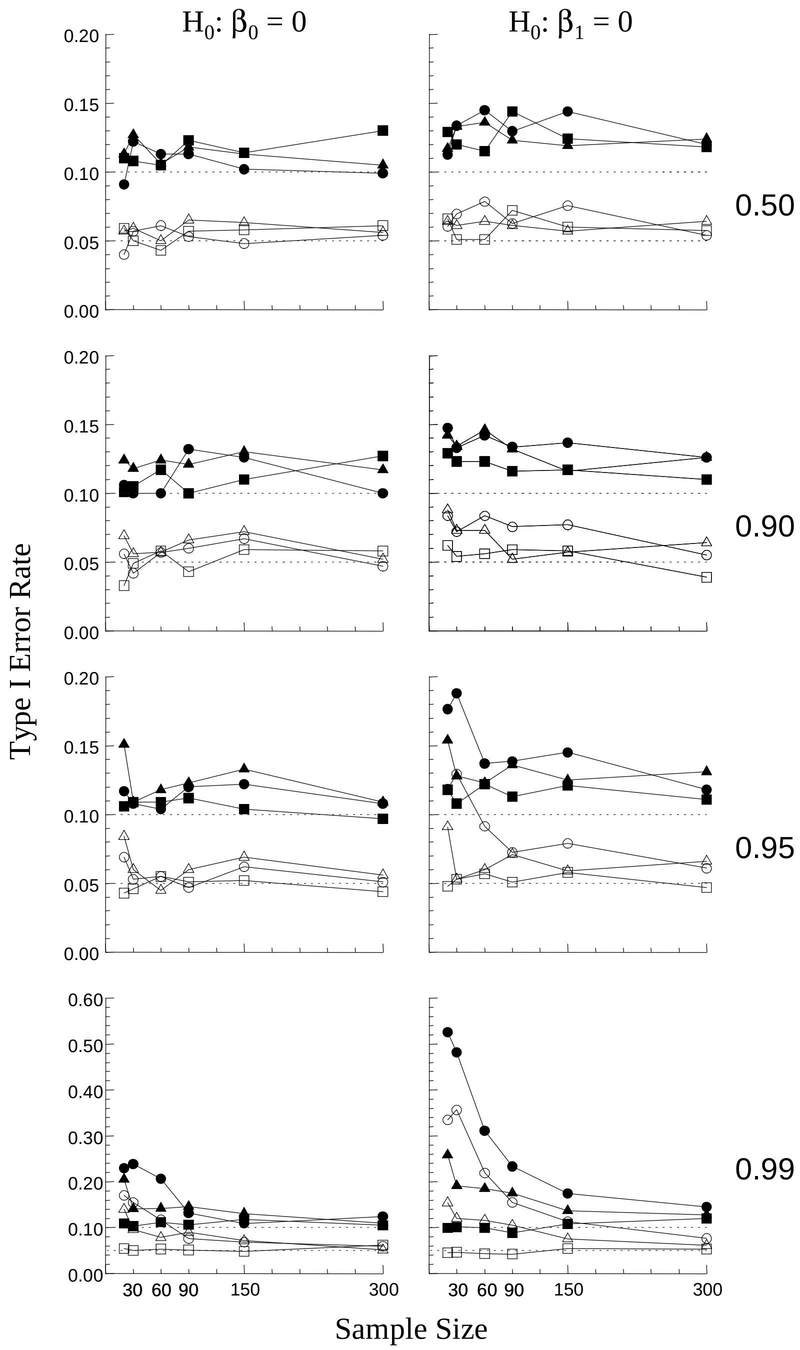


Figure 3.4. Estimated type I error rates for $\alpha = 0.05$ (open) and 0.10 (solid); for the permutation D test for heterogeneous lognormal (circles), normal (triangles), and uniform (squares) error distributions; for H_0: $\beta_0 = 0$ and H_0: $\beta_1 = 0$ in the weighted model $wy = w\beta_0 + w\beta_1 X_1 + w(1 + \gamma X_1)\varepsilon$ with $\gamma = 0.05$ and $w = (1 + \gamma X_1)^{-1}$; for 0.50, 0.90, 0.95, and 0.99 quantiles; and for n = 20, 30, 60, 90, 150, and 300. 1,000 random samples were used at each combination of H_0:, n, and quantile.

discrepancy is similar to that observed when testing the intercept term. The null model implied by the weighted model is forced through the origin (because the column vector of 1's for the intercept have been multiplied by w) and the residuals no longer have their expected properties when the objective function (1) is minimized.

The H_0: $\beta_0 = 0$ also was evaluated in the 2 parameter weighted regression model with heterogeneous errors, $wy = w\beta_0 + w\beta_1 X_1 + w(1 + \gamma X_1)\varepsilon$, with $\gamma = 0.05$ using the known weights $w = (1 + 0.05X_1)^{-1}$, $\beta_1 = 0.10$, and $\beta_0 = 0.0$, 0.5, 1.0, 2.0, and 3.0. Type I error rates were slightly liberal (Fig. 3.4), but no more so than when testing this hypothesis for homogeneous error distributions (Fig. 3.1).

Power to detect $\beta_1 = 0.01$, 0.05, 0.10, and 0.20 for the weighted regression model with heterogeneous errors declined with increasing quantile and decreasing sample size more for the lognormal (Fig. 3.5) than the normal and uniform error distributions (Appendices 3.3 and 3.4). Power for the lognormal error distribution and the 0.99 quantile was unreliable for $n < 150$ because of excessively liberal Type I error rates for smaller sample sizes. Power to detect $\beta_0 = 0.5$, 1.0, 2.0, and 3.0 followed a similar decline with increasing quantile and decreasing samples size as for homogeneous error distributions, becoming almost nonexistent for the 0.99 quantile of the lognormal error distribution (Fig. 3.5). Uniform and normal error distributions had effective power for the 0.99 quantile (Appendices 3.3 and 3.4). Power estimates were assumed to be slightly inflated because Type I error rates were slightly liberal for the weighted regression models.

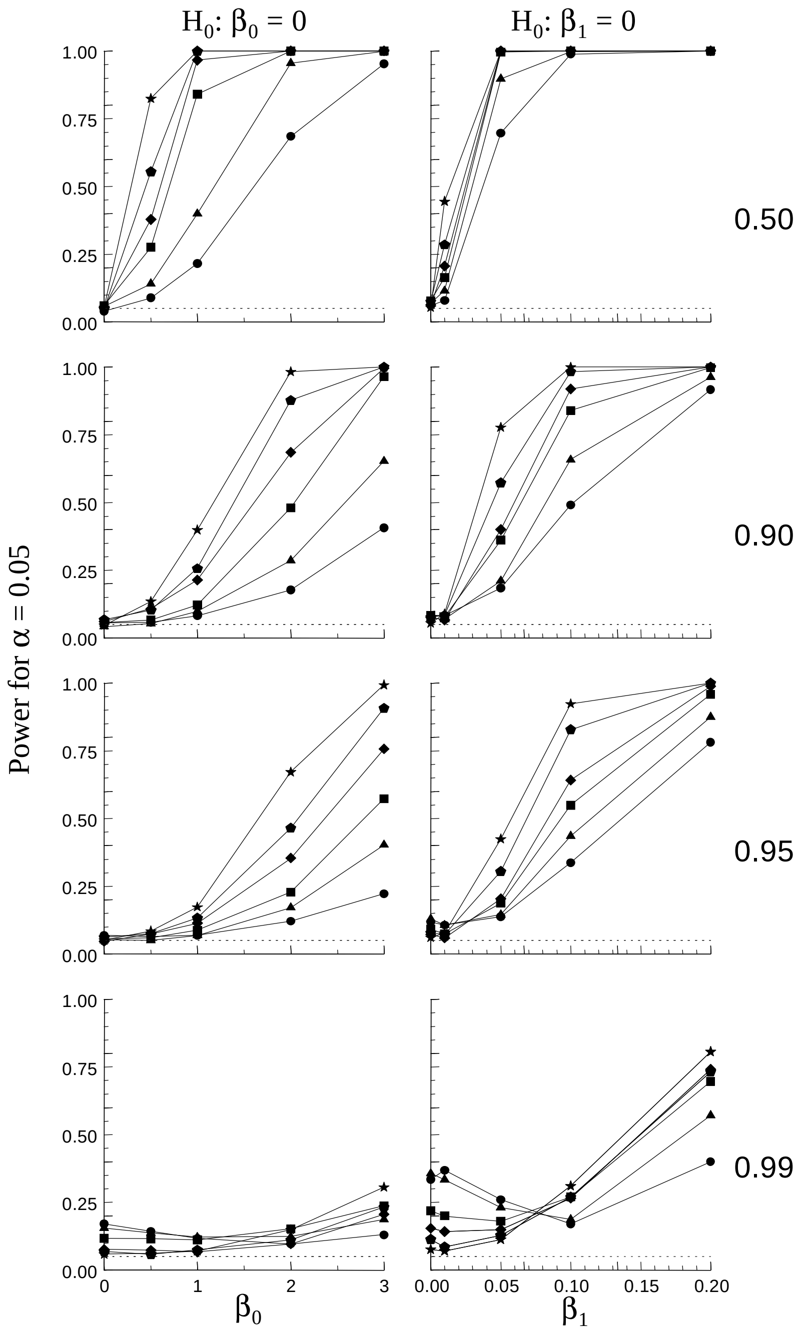


Figure 3.5. Estimated power for $\alpha = 0.05$ for the permutation D test for heterogeneous lognormal error distributions for H_0: $\beta_0 = 0$ and H_0: $\beta_1 = 0$ in the weighted model $wy = w\beta_0 + w\beta_1 X_1 + w(1 + \gamma X_1)\varepsilon$ with $\gamma = 0.05$ and $w = (1 + \gamma X_1)^{-1}$; for $\beta_0 = 0.0$, 0.5, 1.0, 2.0, and 3.0 and for $\beta_1 = 0.0$, 0.01, 0.05, 0.10, and 0.20; for 0.50, 0.90, 0.95, and 0.99 quantiles; and for $n = 20$ (circle), 30 (triangle), 60 (square), 90 (diamond), 150 (pentagon), and 300 (star). 1,000 random samples were used at each combination of effect size, n, and quantile.

4.4 Heterogenous Error Structure - Multiple Regression

The 6 parameter model, $wy = w(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + (1 + \gamma X_1)\varepsilon)$, with $\gamma = 0.05$ and known weights $w = (1 + 0.05X_1)^{-1}$ was evaluated for the full model hypothesis H_0: $\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0$ for β_0 fixed at 36.0 and $\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0$ for Type I error rates, and with $\beta_3 = 0.10, 0.15, 0.20, 0.25$ for power. Type I error rates were slightly liberal for 0.50 - 0.95 quantiles, becoming more liberal with decreasing sample size for the 0.99 quantile, especially for lognormal error distributions (Fig. 3.6). Power estimated with 1 of the 5 slope parameters (β_3) allowed to be nonzero was low to nonexistent for the 0.99 quantile (Fig. 3.8). Power for this and other hypotheses for the multiple regression models was evaluated only for the lognormal error distribution to reduce the amount of computing and reporting. Power for normal and uniform error distributions would be greater than or equal to that for the lognormal error distribution for the quantiles considered.

Type I error rates for subhypotheses involving continuous variables in the 6 parameter weighted model were evaluated for H_0: $\beta_3 = 0$ and H_0: $\beta_3 = \beta_5 = 0$ with $\beta_0 = 36.0$, $\beta_1 = 0.10$, $\beta_2 = -0.005$, $\beta_4 = 2.0$, and $\beta_3 = \beta_5 = 0.0$. As elsewhere, Type I error rates were slightly liberal, becoming more liberal with increasing quantile and decreasing sample size, more so for H_0: $\beta_3 = \beta_5 = 0$ (Fig. 3.7) than for H_0: $\beta_3 = 0$ (Fig. 3.6). Power for H_0: $\beta_3 = 0$ was estimated with $\beta_3 = 0.10, 0.15, 0.20$, and 0.25 for the lognormal error distribution. Power was low for the 0.90 to nonexistent for the 0.95 and 0.99 quantiles (Fig. 3.8).

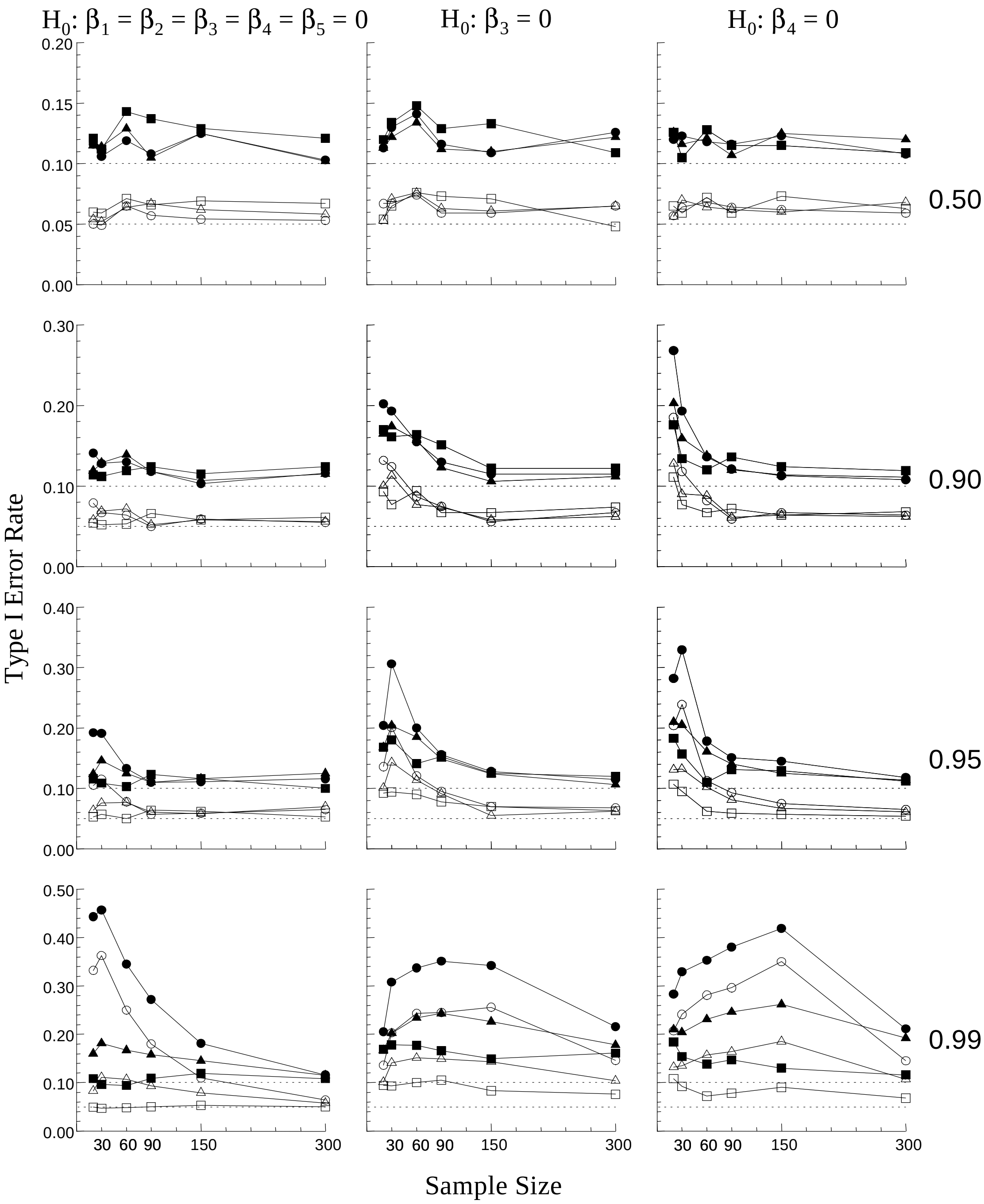


Figure 3.6. Estimated type I error rates for $\alpha = 0.05$ (open) and 0.10 (solid) for the permutation D test for H_0: $\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0$, H_0: $\beta_3 = 0$, and H_0: $\beta_4 = 0$; for heterogeneous lognormal (circles), normal (triangles) and uniform (squares) error distributions with $\gamma = 0.05$ in the weighted model $wy = w(\beta_0 + \beta_1X_1 + \beta_2X_2 + \beta_3X_3 + \beta_4X_4 + \beta_5X_5 + (1 + \gamma X_1)\varepsilon)$ with $w = (1 + \gamma X_1)^{-1}$; for 0.50, 0.90, 0.95, and 0.99 quantiles; and for n = 20, 30, 60, 90, 150, and 300. 1,000 random samples were used at each combination of H_0, n, and quantile.

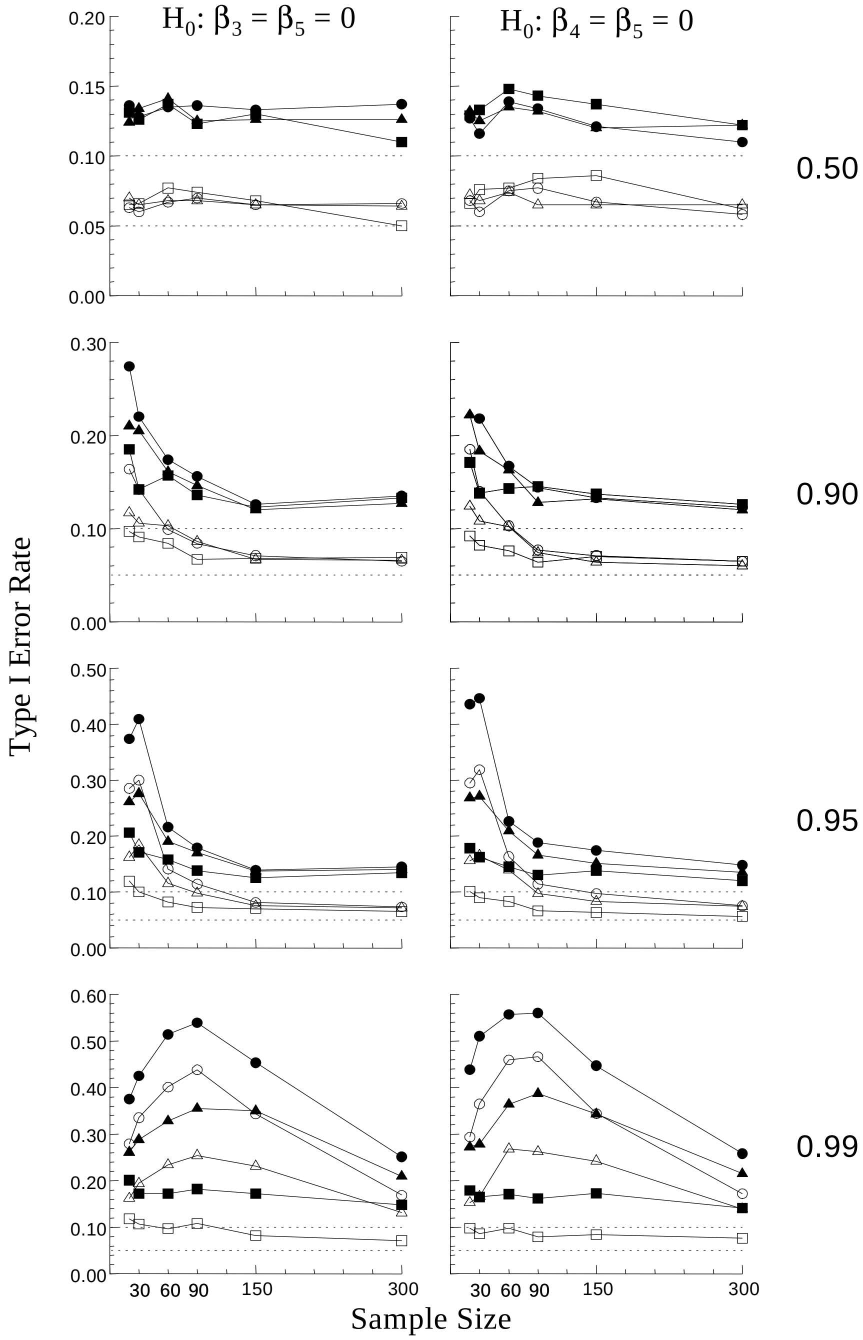


Figure 3.7. Estimated type I error rates for $\alpha = 0.05$ (open) and 0.10 (solid) for the permutation D test for H_0: $\beta_3 = \beta_5 = 0$ and H_0: $\beta_4 = \beta_5 = 0$; for heterogeneous lognormal (circles), normal (triangles) and uniform (squares) error distributions with $\gamma = 0.05$ in the weighted model $wy = w(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + (1 + \gamma X_1)\varepsilon)$ with $w = (1 + \gamma X_1)^{-1}$; for 0.50, 0.90, 0.95, and 0.99 quantiles; and for n = 20, 30, 60, 90, 150, and 300. 1,000 random samples were used at each combination of H_0, n, and quantile.

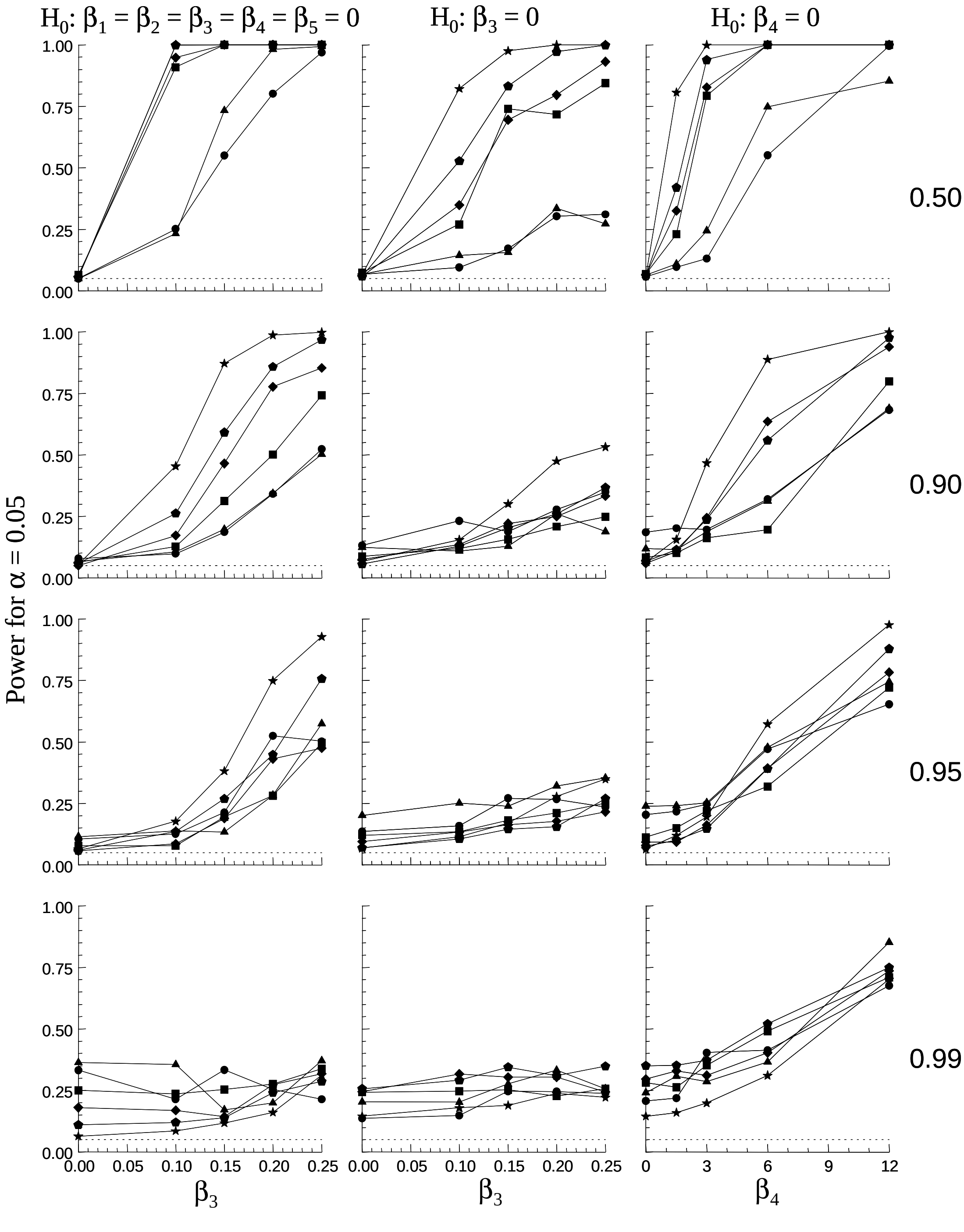


Figure 3.8. Estimated power for $\alpha = 0.05$ for the permutation *D* tests for H_0: $\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0$, H_0: $\beta_3 = 0$, and H_0: $\beta_4 = 0$; for heterogeneous lognormal error distributions with $\gamma = 0.05$ in the weighted model $wy = w(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + (1 + \gamma X_1)\varepsilon)$; for 0.50, 0.90, 0.95, and 0.99 quantiles; and for n = 20 (circle), 30 (triangle), 60 (square), 90 (diamond),150 (pentagon), and 300 (star). 1,000 random samples were used at each combination of H_0, n, and quantile.

Subhypotheses involving categorical predictors in the 6 parameter weighted model were evaluated for H_0: $\beta_4 = 0$ and H_0: $\beta_4 = \beta_5 = 0$ with $\beta_0 = 36.0$, $\beta_1 = 0.10$, $\beta_2 = -0.005$, $\beta_3 = 0.05$, and $\beta_4 = \beta_5 = 0.0$. Type I error rates for H_0: $\beta_4 = 0$ (Fig. 3.6) and H_0: $\beta_4 = \beta_5 = 0$ (Fig. 3.7) had similar patterns as the Type I error rates for subhypotheses for continuous predictors evaluated above. Power was evaluated for the subhypothesis H_0: $\beta_4 = 0$ for $\beta_4 = 1.5$, 3.0, 6.0, and 12.0 and the lognormal error distribution. Power declined with increasing quantiles and decreasing sample size but was grossly inflated at smaller sample sizes for 0.95 and 0.99 quantiles (Fig. 3.8) because of excessively liberal Type I error rates.

5. Example Application

I constructed confidence intervals for quantile regression estimates of Lahontan cutthroat trout *Oncorhynchus clarki henshawi* density (trout m^{-1}) as a function of stream channel morphology (width:depth ratio) for 13 small streams in Nevada sampled over 7 years (Dunham et al. 2002). Width:depth ratio is a measure that integrates stream channel characteristics thought to be related to small stream integrity and, thus, fish populations and is easily measured for assessing fish habitat conditions and land use impacts over large regions. Lahontan cutthroat trout are a threatened species of special interest to federal land management agencies.

Here I considered the nonlinear model $y = \exp(\beta_0 + \beta_1 X_1 + \varepsilon)$, where y is trout m^{-1} and X_1 is width:depth ratio, for $n = 71$ observations of streams for 1993 to 1999 (Dunham et al. 2002). The model was estimated in the weighted linear form $\ln wy = w\beta_0 + w\beta_1 X_1 + w\varepsilon$ and estimates for selected regression quantiles were plotted by exponentiating to back transform to the nonlinear form (Fig. 3.9). The vector of weights w were identical to those used with the quantile rankscore tests (Chapter 2). Weights were estimated by computing the average pairwise differences between the 76 unweighted regression quantile estimates $b_0(\tau)$ to estimate γ_0 and $b_1(\tau)$ to estimate γ_1 in the standard deviation function $\gamma_0 - \gamma_1 X_1$ and then taking the reciprocal, $w = (1.310 - 0.017X_1)^{-1}$. Estimates of parameters for all quantiles were plotted as a step function with 90% confidence intervals for 19 quantiles between 0.05 and 0.95 by increments of 0.05 (Fig. 3.9).

Interval endpoints were estimated by inverting the D test as an alternative to inverting the quantile rankscore tests used by Dunham et al. (2002) and in Chapter 2. Starting values for the manual iteration of the test inversion were based on the interval endpoints estimated by the rankscore tests (Chapter 2). These values were then used as hypothesized parameter values of $\xi(\tau)$ in the transformation $\mathbf{y} - \mathbf{X_2}\xi(\tau)$ to test the H_0: $\boldsymbol{\beta}_2(\tau) = \xi(\tau)$ with (3), where $\boldsymbol{\beta}_2$ was either β_0 or β_1 depending on the parameter being tested. I used $m + 1 = 100{,}000$ permutations to compute probabilities for the D tests associated with confidence interval endpoints.

The 90% confidence intervals estimated by inverting the D test (Fig. 3.9) were smoother than those for the rankscore tests (Chapter 2) because the permutation distribution for the D test was much more continuous than the distributions of the rankscore tests. Linear interpolation between hypothesize parameter values was not required to achieve $P = \alpha = 0.10$ with the D test as it was for the rankscore tests (Koenker 1994, Chapter 2). The D test based 90% confidence intervals were slightly

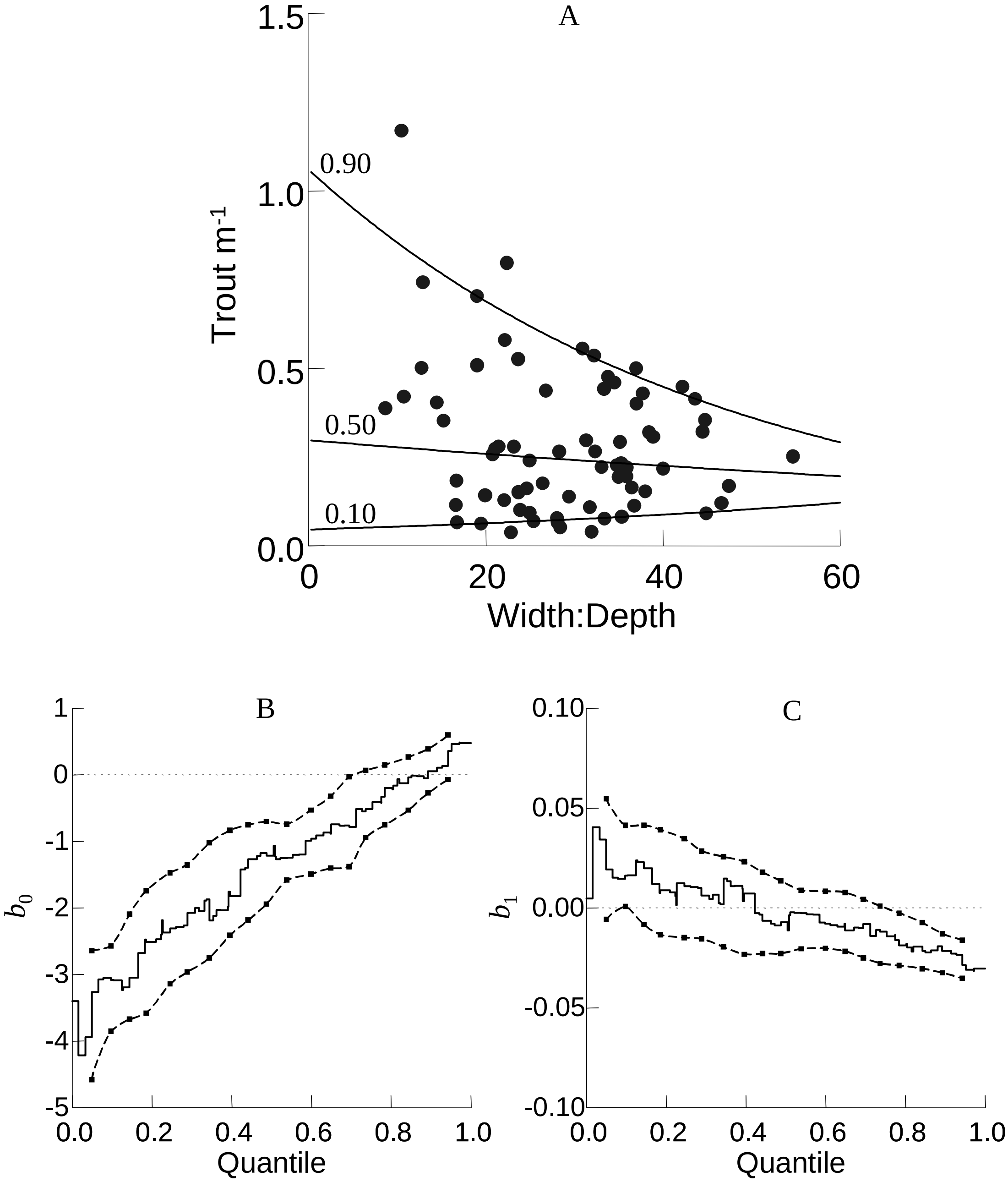


Figure 3.9. (A) Lahontan cutthroat trout m^{-1} and width:depth ratios for 13 small streams sampled 1993-1999 (n = 71); exponentiated estimates for 0.90, 0.50, and 0.10 regression quantiles for the weighted model $(\ln y)w = (\beta_0 + \beta_1 X_1 + (\gamma_0 - \gamma_1 X_1)\varepsilon)w$, $w = (1.310 - 0.017X_1)^{-1}$. Step functions (solid lines) for estimates of β_0 (B) and β_1 (C) by quantiles [0, 1] are bracketed by pointwise 90% confidence intervals (dashed lines) based on inverting the D permutation test.

narrower for lower quantiles than those based on inverting rankscore tests (Chapter 2), but nearly identical in width for the upper quantiles. Differences between the *D* test and the quantile rankscore based intervals (Chapter 2) were not sufficient to alter any conclusions about the effects of width:depth ratios on cutthroat trout populations. Confidence bands estimated by the *D* test supported an interpretation that increasing stream width:depth ratios from 15 to 45 decreased the highest 20% of trout densities ($\tau \geq 0.80$) by 9 to 65% [exp(-0.003 × 30) = 0.914 and exp(-0.035 × 30) = 0.350], similar to conclusion based on the quantile rankscore test intervals (Dunham et al. 2002, Chapter 2).

6. Discussion

Although the drop in dispersion permutation *D* test had better power than the quantile rankscore tests for hypotheses where both maintained reasonable Type I errors, it had extremely liberal Type I error rates at smaller samples and less extreme quantiles than the quantile rankscore tests (Chapter 2). Type I error rate failure occurred more rapidly with decreasing sample size and increasing quantile for the lognormal error distribution than for normal or uniform error distributions because it has a long-tail with low density of observations. My example application with the Lahontan cutthroat trout data suggested that the differences between the drop in dispersion permutation test and quantile rankscore tests may not always be sufficient to substantively affect the interpretation of an analysis when quantiles used are not too extreme (e.g., $0.05 \leq \tau \leq 0.95$). When estimating models for more extreme quantiles (e.g., $\tau = 0.99$), fairly large samples ($n > 300$) will be required for models with more than just a few parameters to ensure reliable confidence intervals based on either class of test.

The slightly liberal nature of the permutation *D* test when testing the intercept term in unweighted models and any parameter in weighted models was consistent with simulation results for permutation versions of the rankscore tests (Chapter 2). There is additional sampling variation not accounted for by the permutation distribution of the test statistics when the null model was constrained through the origin. If the number of positive, negative, and zero residuals are denoted by N^+, N^-, N^0, respectively, and if $N^0 = p - q$ under the null model, then there are at most $n\tau$ negative residuals ($N^- \leq n\tau \leq N^- + N^0$) and at most $n(1 - \tau)$ positive residuals ($N^+ \leq n[1 - \tau] \leq N^+ + N^0$) when the null model includes an intercept (Koenker and Bassett 1978, Koenker and Portnoy 1996). When the null model does not include an intercept, the limits on the number of positive (negative) residuals exceeded these values by amounts consistent with binomial random variation with success probability $1 - \tau$ (or τ for negative residuals). This is similar to least squares regression models forced through the origin which do not have the mean of the residuals equal to zero. Legendre and Desdevises (In Press) proposed a solution for permutation tests for least squares regression by using a double permutation scheme where the first step varies the number of positive (negative) residuals as a binomial random variable with success probability 0.5. and the second step permutes these residuals across the rows of **X**. This procedure was easily modified for the permutation version of the quantile rankscore test (Chapter 2) but needs to be investigated for application to the *D* test.

My simulation experiment avoided the issue of how to estimate weights for heteroscedastic models by using the known standard deviation function. In applications, this function is not known and the weights must be estimated. I used a simple pairwise difference approach based on the initial unweighted estimates for estimating weights in my example application. Other approaches for estimating weights include regressing absolute values of residuals from an unweighted fit of the 0.5 quantile on the independent variables for linear location-scale models (Zhou and Portnoy 1998) and the sparsity estimation approach for more general heteroscedastic models (Koenker and Machado 1999).

Acknowledgments

J. D. Richards provided programming support for the simulations. M. J. Anderson, P. Good, R. Koenker, and J. E. Roelle reviewed earlier drafts of the manuscript. J. B. Dunham provided the Lahontan cutthroat trout data.

Literature Cited

Allen, A. W., B. S. Cade, and M. W. Vandever. 2001. Effects of emergency haying on vegetative characteristics within selected conservation reserve program fields in the northern Great Plains. J. Soil and Water Conservation **56**: 120-125.

Anderson, M. J., and P. Legendre. 1999. An empirical comparison of permutation methods for tests of partial regression coefficients in a linear model. Journal Statistical Computation and Simulation **62**:271-303.

Anderson, M. J., and J. Robinson. 2001. Permutation tests for linear models. Australian New Zealand Journal of Statistics **43**:75-88.

Barrodale, I., and F. D. K. Roberts. 1974. Algorithm 478: Solution of an overdetermined system of equations in the l_1 norm. Communications of the Association for Computing Machinery **17**:319-320.

Buchinsky, M. 1991. The theory and practice of quantile regression. Ph.D dissertation, Harvard University, Cambridge Mass. USA. 211pp.

Cade, B. S., and Q. Guo. 2000. Estimating effects of constraints on plant performance with regression quantiles. Oikos **91**: 245-254.

Cade, B. S., and J. D. Richards. 1996. Permutation tests for least absolute deviation regression. Biometrics **52**:886-902.

Cade, B. S., J. W. Terrell, and R. L. Schroeder. 1999. Estimating effects of limiting factors with regression quantiles. Ecology **80**:311-323.

Chernozhukov, V., and L. Umantsev. 2001. Conditional value-at-risk: Aspects of modeling and estimation. Empirical Economics **26**: 271-292.

Commenges, D. In Press. Transformations which preserve exchangeability and application to permutation tests. Journal of Nonparametric Statistics.

Dunham, J. B., B. S. Cade, and J. W. Terrell. 2002. Influences of spatial and temporal variation on fish-habitat relationships defined by regression quantiles. Transactions of the American Fisheries Society **131**: 86-98.

Eastwood, P. D., G. J. Meaden, and A. Grioche. 2001. Modeling spatial variations in spawning habitat suitability for the sole *Solea solea* using regression quantiles and GIS procedures. Marine Ecology Progress Series **224**: 251-266.

Freedman, D., and D. Lane. 1983. A nonstochastic interpretation of reported significance levels. Journal of Business and Economic Statistics **1**:292-298.

Gutenbrunner, C., and J. Jurečková. 1992. Regression rank scores and regression quantiles. Annals of Statisitics **20**: 305-330.

Gutenbrunner, C., J. Jurečková, R. Koenker, and S. Portnoy. 1993. Tests of linear hypotheses based on regression rank scores. Nonparametric Statistics **2**:307-331.

Haire, S. L., C. E. Bock, B. S. Cade, and B. C. Bennett. 2000. The role of landscape and habitat characteristics in limiting abundance of grassland nesting songbirds in an urban open space. Landscape and Urban Planning **48**(1-2):65-82.

Huston, M. A. 2002. Introductory essay: Critical issues for improving predictions. Pages 7-21 *in* J. M. Scott, P. J. Heglund, and M. L. Morrison, editors. Predicting species occurrences: issues of accuracy and scale. Island Press, Covelo, California, USA.

Kennedy, P. E., and B. S. Cade. 1996. Randomization tests for multiple regression. Communications in Statistics - Simulation and Computation **25**: 923-936.

Koenker, R. 1987. A comparison of asymptotic testing methods for l_1-regression. Pages 287-295 *in* Y. Dodge, editor. Statistical data analysis based on the L_1-norm and related methods. Elsevier Science Publishers B. V. (North-Holland), Amsterdam.

Koenker, R. 1994. Confidence intervals for regression quantiles. *Pages* 349-359 *in* P. Mandl and M. Hušková, editors. Asymptotic statistics: Proceedings of the 5th

Prague Symposium. Physica-Verlag

Koenker, R. 1997. Rank tests for linear models. Pages 175-199 *in* G. S. Maddala and C. R. Rao, editors. Handbook of Statistics, vol 15.

Koenker, R., and G. Bassett. 1978. Regression quantiles. Econometrica **46**:33-50.

Koenker, R., and G. Bassett. 1982. Robust tests for heteroscedasticity based on regression quantiles. Econometrica **50**: 43-61.

Koenker, R. and V. d'Orey. 1987. Computing regression quantiles. Applied Statistics **36**:383-393.

Koenker, R. and V. d'Orey. 1994. A remark on algorithm AS229: computing dual regression quantiles and regression rank scores. Applied Statistics **43**:410-414.

Koenker, R., and J. A. F. Machado. 1999. Goodness of fit and related inference processes for quantile regression. Journal of the American Statistical Association **94**:1296-1310.

Koenker, R., and S. Portnoy. 1996. Quantile regression. University of Illinois at Urbana-Champaign, College of Commerce and Business Administration, Office of Research Working Paper 97-0100. 77pp.

Koenker, R., and Q. Zhao. 1994. *L*-estimation for linear heteroscedastic models. Nonparametric Statistics **3**:223-235.

Legendre, P. 2000. Comparison of permutation methods for the partial correlation and partial Mantel tests. Journal of Statistical Computation and Simulation **67**:37-73.

Legendre, P., and Y. Desdevises. (In press). Independent contrasts and regression through the origin. Systematic Biology.

Randles, R. H. 1984. On tests applied to residuals. Journal of the American Statistical Association **79**:349-354.

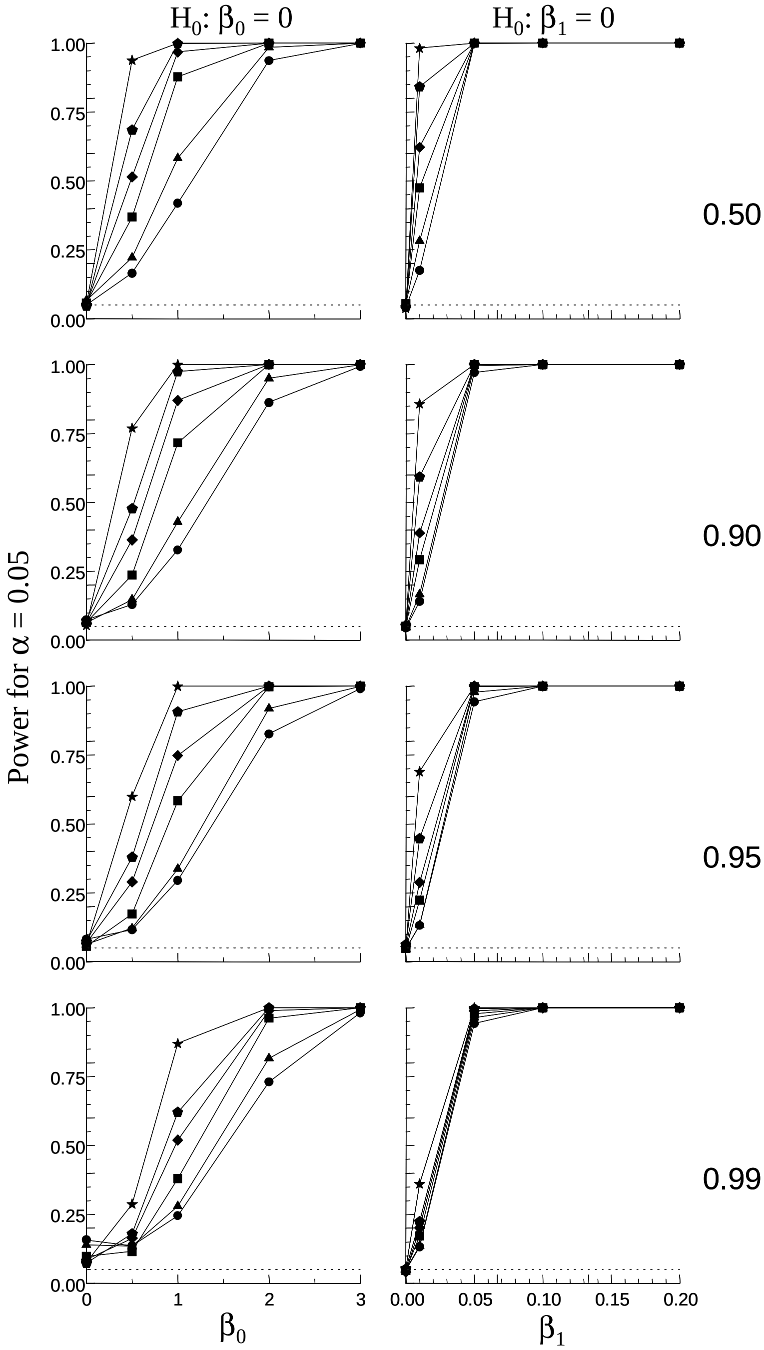
Scharf, F. S., F. Juanes, and M. Sutherland. 1998. Inferring ecological relationships from the edges of scatter diagrams: comparison of regression techniques. Ecology **79**:448-460.

Terrell, J. W., B. S. Cade, J. Carpenter, and J. M. Thompson. 1996. Modeling stream fish habitat limitations from wedged-shaped patterns of variation in standing stock. Transactions of the American Fisheries Society **125**:104-117.

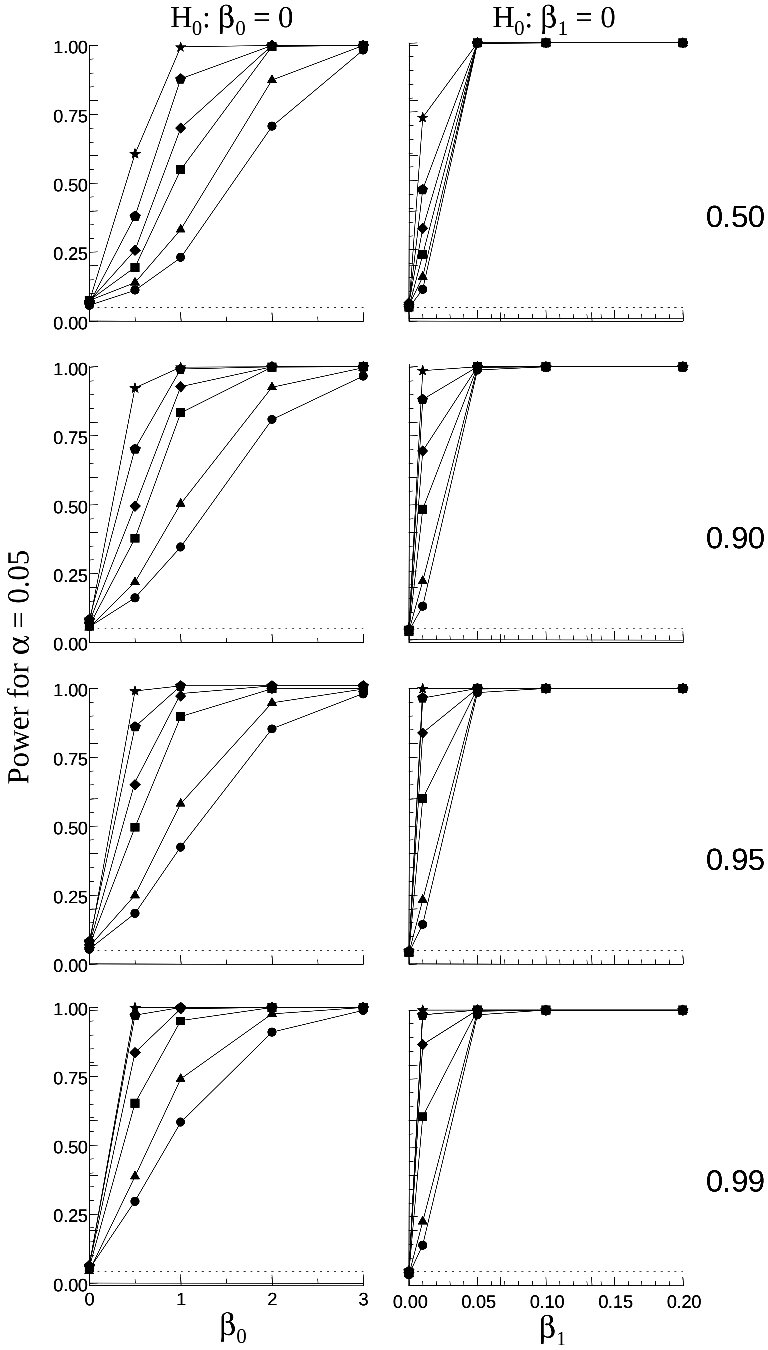
Zhou, K. G., and S. L. Portnoy. 1998. Statistical inference on heteroscedastic models based on regression quantiles. Nonparametric Statistics **9**:239-260.

Appendix 3

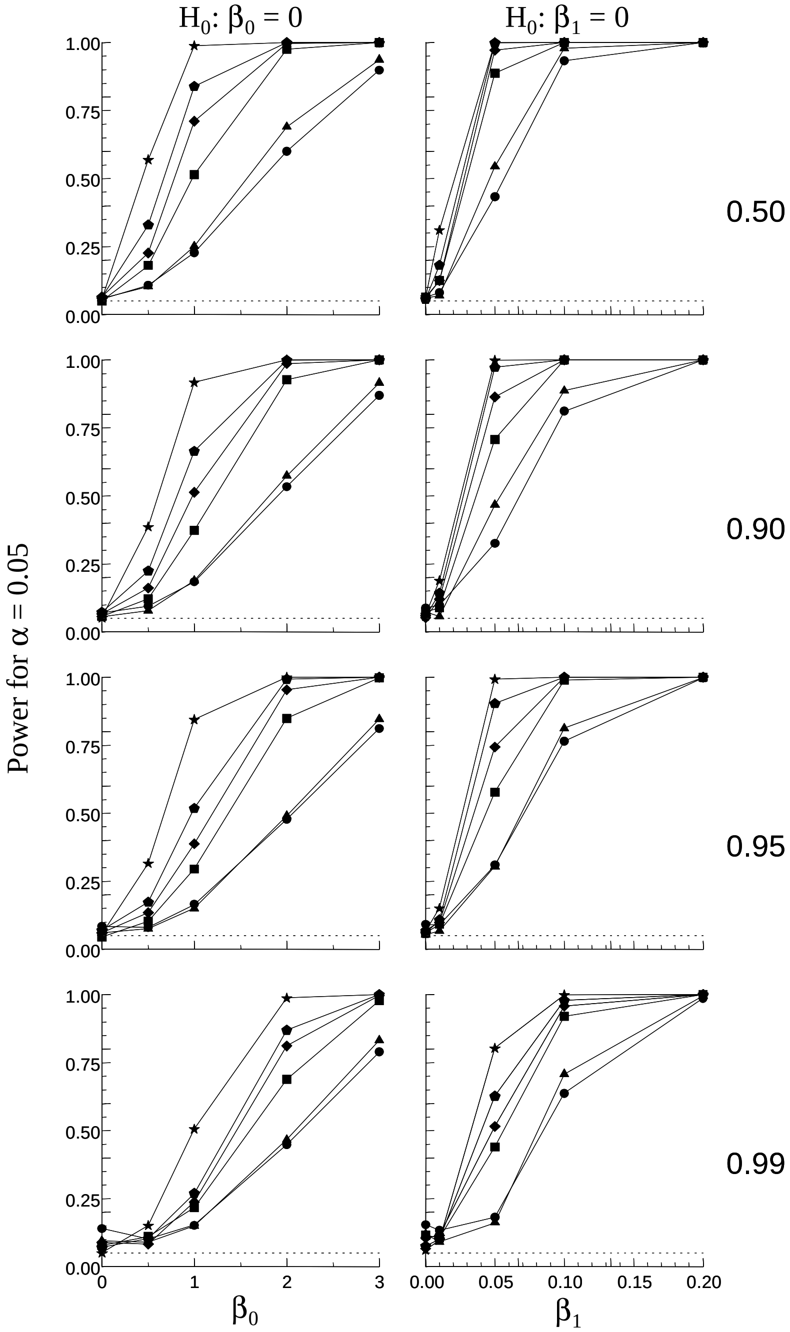
Simulation Results for Normal and Uniform Error Distributions



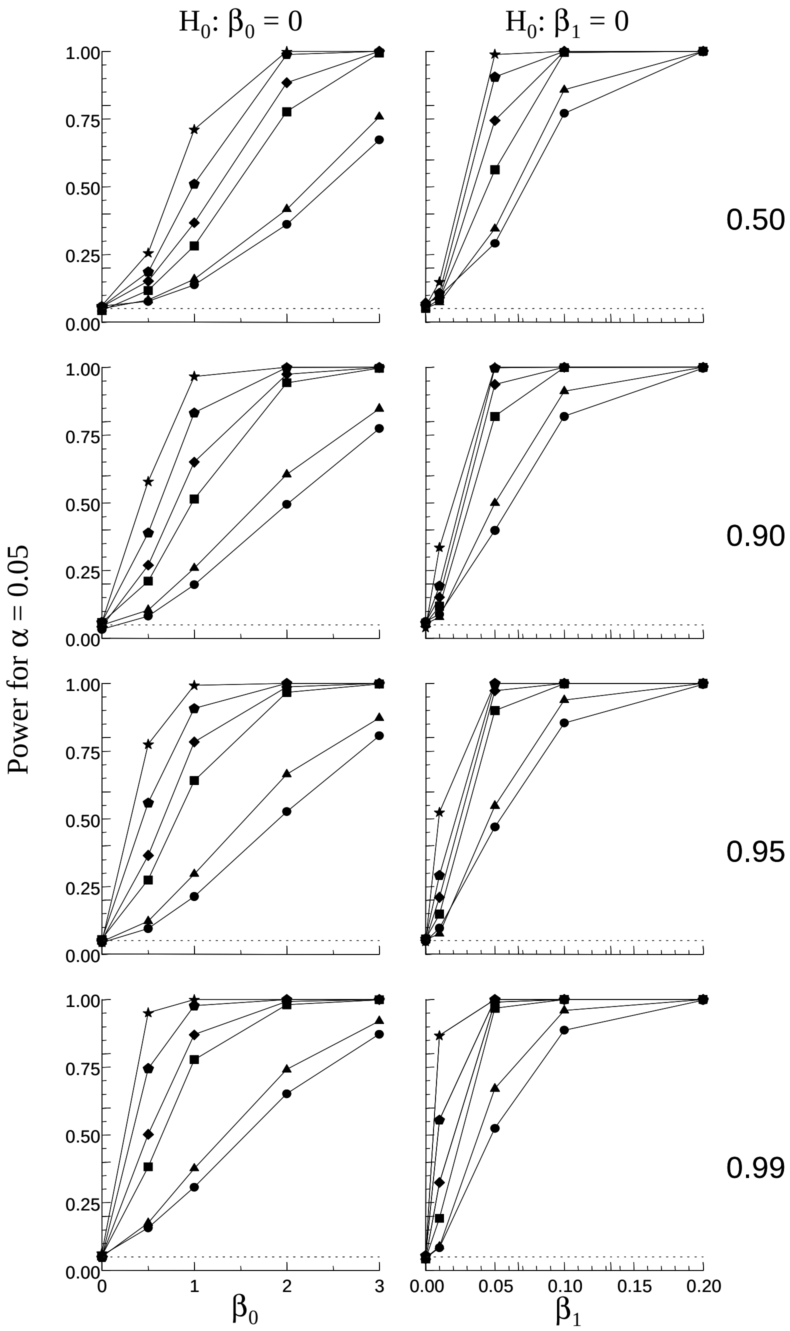
Appendix 3.1. Estimated power for $\alpha = 0.05$ for the permutation D tests for homogeneous normal error distributions for H_0: $\beta_0 = 0$ and H_0: $\beta_1 = 0$ in the model $y = \beta_0 + \beta_1 X_1 + \varepsilon$; for β_0 = 0.0, 0.5, 1.0, 2.0, and 3.0 and for β_1 = 0.0, 0.01, 0.05, 0.10, and 0.20; for 0.50, 0.90, 0.95, and 0.99 quantiles; and for n = 20 (circle), 30 (triangle), 60 (square), 90 (diamond), 150 (pentagon), and 300 (star). 1,000 random samples were used at each combination of effect size, n, and quantile.



Appendix 3.2. Estimated power for $\alpha = 0.05$ for the permutation D tests for homogeneous uniform error distributions for H_0: $\beta_0 = 0$ and H_0: $\beta_1 = 0$ in the model $y = \beta_0 + \beta_1 X_1 + \varepsilon$; for $\beta_0 = 0.0, 0.5, 1.0, 2.0$, and 3.0 and for $\beta_1 = 0.0, 0.01, 0.05, 0.10$, and 0.20; for 0.50, 0.90, 0.95, and 0.99 quantiles; and for $n = 20$ (circle), 30 (triangle), 60 (square), 90 (diamond), 150 (pentagon), and 300 (star). 1,000 random samples were used at each combination of effect size, n, and quantile.



Appendix 3.3. Estimated power for $\alpha = 0.05$ for the permutation D test for heterogeneous normal error distributions for H_0: $\beta_0 = 0$ and H_0: $\beta_1 = 0$ in the weighted model $wy = w\beta_0 + w\beta_1 X_1 + w(1 + \gamma X_1)\varepsilon$ with $\gamma = 0.05$ and $w = (1 + \gamma X_1)^{-1}$; for $\beta_0 = 0.0$, 0.5, 1.0, 2.0, and 3.0 and for $\beta_1 = 0.0$, 0.01, 0.05, 0.10, and 0.20; for 0.50, 0.90, 0.95, and 0.99 quantiles; and for $n = 20$ (circle), 30 (triangle), 60 (square), 90 (diamond), 150 (pentagon), and 300 (star). 1,000 random samples were used at each combination of effect size, n, and quantile.



Appendix 3.4. Estimated power for $\alpha = 0.05$ for the permutation D test for heterogeneous lognormal error distributions for H_0: $\beta_0 = 0$ and H_0: $\beta_1 = 0$ in the weighted model $wy = w\beta_0 + w\beta_1X_1 + w(1 + \gamma X_1)\varepsilon$ with $\gamma = 0.05$ and $w = (1 + \gamma X_1)^{-1}$; for $\beta_0 = 0.0$, 0.5, 1.0, 2.0, and 3.0 and for $\beta_1 = 0.0$, 0.01, 0.05, 0.10, and 0.20; for 0.50, 0.90, 0.95, and 0.99 quantiles; and for $n = 20$ (circle), 30 (triangle), 60 (square), 90 (diamond), 150 (pentagon), and 300 (star). 1,000 random samples were used at each combination of effect size, n, and quantile.

Chapter 4

Estimating Effects of Limiting Habitat Resources: Hidden Bias and Spatial Structure

Abstract: Simulations from a large (N = 10,000) finite population representing grid areas on a landscape were made to demonstrate various forms of hidden bias that might occur when the effect of a measured habitat variable on some animal was confounded with the effect of another unmeasured variable (spatially and not spatially structured). Regression quantile ($0 \leq \tau \leq 1$) parameters for linear models that excluded the important, unmeasured variable were used to evaluate bias relative to parameters from the generating model. Depending on whether interactions of the measured habitat and unmeasured variable were negative (interference interactions) or positive (facilitation interactions), either upper ($\tau > 0.5$) or lower ($\tau < 0.5$) quantile regression parameters were less biased than mean rate parameters. Sampling (n = 20 - 300) simulations demonstrated that regression quantile estimates and confidence intervals constructed by inverting rankscore tests provided valid coverage of these biased parameters. Local forms of quantile weighting were required for obtaining correct Type I error rates and confidence interval coverage. Heterogeneous, nonlinear response patterns occurred in simulations with correlations between the measured and unmeasured variables. When the unmeasured variable was spatially structured, variation in parameters across quantiles associated with effects of the habitat variable were reduced by modeling the spatial trend surface as a cubic polynomial of location coordinates, but substantial

hidden bias in the parameters remained. Quantile regression was used to estimate effects of physical habitat resources on a bivalve mussel (*Macomona liliana*) in the spatially structured landscape of a New Zealand harbor.

1. Introduction

The relationship between an organism and its habitat is of theoretical interest in ecology because it is fundamentally tied to questions about distribution and abundance (Wiens 1989, Huston 2002). Habitat relationships also are important in natural resource management because environmental regulations in the United States (e.g., National Environmental Policy Act, Fish and Wildlife Coordination Act, National Forest Management Act) mandate that management agencies consider impacts to fish and wildlife habitat in their land use planning (Morrison et al. 1998). Mathematical and statistical models commonly are used for quantifying the relationship between an organism and the resources provided by its habitat. Habitat models are used for predicting changes in distribution and abundance due to changes in resources driven by alternative land management or environmental changes (Stauffer 2002). Reliability of quantitative predictions from animal habitat models has been questioned, however, because factors other than the resources provided by habitat may limit animal populations (Rotenberry 1986, Fausch et al. 1988, Terrell et al. 1996, Terrell and Carpenter 1997). Typically, not all factors that limit populations are measured and included in the models, either due to logistical constraints or because they are unknown. As a consequence, predicted responses to changes in habitat often lack the generality to be considered reliable statements of outcomes likely to occur at other times or places than those originally sampled. This hinders both the development of general theory related to resource selection and the utility of models for predicting outcomes of alternative management or conservation actions.

We can envision the distribution and abundance of any species as being constrained by biophysical factors (e.g., climate, soil productivity), habitat resources (e.g., vegetation providing food and cover), and interspecific (e.g., competition and predation) and intraspecific (e.g., density dependent behavioral responses) biotic interactions (Morrison 2001, Huston 2002, O'Connor 2002). When none of the factors are limiting over some interval of time and space, then the species will be locally abundant. When any single factor is limiting, the species will be constrained to a locally lower abundance than expected if all factors are permissive. Processes associated with the constraints operate at different rates, slower for most biophysical factors and faster for biotic interactions. If factors that are the active constraint limiting species abundance at some sample locations are unmeasured, then the species response may exhibit heterogeneous variation across levels of the measured habitat resources simply because they are not limiting at all times or locations sampled (Kaiser et al. 1994, Cade et al. 1999). Heterogeneity arises due to interactions among the multiple biotic and abiotic factors that affect growth, survival, and reproduction of an organism, where the factor that is limiting differs among sample locations and times (Van Horne and Wiens 1991, Huston 2002). When we measure only a subset of the potential limiting factors such as habitat resources, it is reasonable to expect a rather large component of unexplained variation to remain in our models, especially as we increase the spatial and temporal extent of our sampling. The variance amplification hypothesis of Huston (2002) suggests that variation in abundance (or other measures such as biomass, survival, or fecundity) increases as habitat quality increases because large variation in population size can only occur where levels of habitat resources permit high abundance when other factors are not limiting.

Huston (2002) and O'Connor (2002) suggested that viewing relationships between organisms and resources provided by their habitat as constraints rather than as correlates is a paradigm shift affecting how we model animal habitat relationships. The essence of this idea is that much of the useful information about how organisms respond to changes in levels of resources may not be found in statistical estimates of rates of change in mean responses but in estimates of rates of change near maximum responses. Changes in responses near the extremes are thought to better represent rates of change when habitat is the constraint rather than other unmeasured processes. Rates of change in a response variable (y) as a function of some predictor variables (X) differ from the center to the extremes of the distributions in heteroscedastic regression models by definition (Terrell et al. 1996, Cade et al. 1999). Statistical difficulties associated with estimating effects at the extremes of heterogeneous response distributions and some solutions have been discussed by Kaiser et al. (1994), Terrell et al. (1996), Thomson et al. (1996), Cade et al. (1999), and Bi et al. (2000).

Quantile regression has been used to estimate effects of ecological limiting factors (Scharf et al. 1998, Cade et al. 1999, Cade and Guo 2000, Huston 2002) because it provides statistical estimates of rates of change in selected or all parts of a response variable distribution. Because quantile regression estimates rates of change across all parts of a response distribution, it is especially informative for modeling heterogeneous distributions like those in animal habitat relationships (Terrell et al. 1996, Cade et al. 1999, Dunham et al. 2002, Huston 2002). It is possible to focus estimation on a selected part of the response distribution near the maximum (e.g, 90^{th} to 99^{th} percentiles) if it is reasonable to assume that the unmeasured processes are only likely to reduce rates of change in the responses (Kaiser et al. 1994, Terrell et al. 1996, Cade et al. 1999, Cade and Guo 2000). This is implicit in the variance amplification hypothesis of Huston (2002). Quantile regression has been used to estimate relationships between stream fish populations and their habitat (Terrell et al. 1996, Dunham et al. 2002); ocean fish and spawning habitat (Eastwood et al. 2001); and breeding grassland birds, habitat, and landscape metrics (Haire et al. 2000).

My objectives were to further explore assumptions about confounded relations between measured habitat variables and unmeasured variables for other important processes and to evaluate the statistical performance of regression quantile estimates and rankscore tests under these conditions of hidden bias. I examined assumptions about the interactions among habitat resources and other limiting factors that would support focusing on estimated rates of change in selected portions of the species response distribution (e.g., the upper quantiles). Sampling distributions for estimates of various quantiles of a species response distribution and associated rankscore test statistics were simulated for a range of interaction and correlation structures between limiting habitat resources, which were considered measured, and some additional limiting factors, which were considered unmeasured. The estimating models contained various degrees of hidden bias (sensu Rosenbaum 1991, 1995, 1999) because rates of change estimated for the measured variables were confounded with effects of the unmeasured variables. Thus, unlike the analyses in Chapters 2 and 3, heterogeneous variance structure of the simulated data was not completely specified by a function of the measured predictors included in the estimating models. Simulated data were generated to mimic spatially structured processes on a landscape. The potential to account for unmeasured limiting factors by modeling spatial trend (Borcard et al. 1992) with quantile regression also was investigated, and a case study was conducted on previously published data (Legendre et al. 1997).

2. Quantile Regression Models With Unmeasured Variables

To explore patterns of heterogeneity due to missing information on some important limiting factor other than habitat resources, I extended the linear model assumptions beyond those considered by Cade et al. (1999) and Huston (2002). Data were generated from a 2 variable linear model with interaction, $y = \theta_0 X_0 + \theta_1 X_1 + \theta_2 X_2 + \theta_3 X_1 X_2 + \varepsilon$, where y was the dependent response variable, X_0 was 1 for the intercept, X_1 (uniform [0, 50]) was the measured habitat variable, X_2 (uniform [0, 4,000]) was a variable for some other limiting process that was not measured and available for the estimating model, and ε was a random error term that was independent and identically distributed (iid). Error distributions were lognormal (median = 0, σ = 0.75) to create asymmetric distributions or uniform [-0.50, 0.50] to create symmetric distributions. By varying the correlation between X_1 and X_2 and direction and size of interaction effects due to θ_3, it was possible to simulate a range of linear, nonlinear, homogeneous, and heterogeneous distribution patterns associated with an estimating model, $y = \beta_0 X_0 + \beta_1 X_1 + \varepsilon'$, where ε' includes the generating error term plus the effect of unmeasured covariates, i.e., $\varepsilon' = \varepsilon + \theta_2 X_2 + \theta_3 X_1 X_2$ (Table 4.1). Note that β_0 and β_1 will not in general be equivalent to θ_0 and θ_1, although they may differ less for some quantiles.

Table 4.1. Parameter values in hidden bias simulations and direction of bias in $\beta_1(\tau)$ relative to θ_1 where generating models were $y = \theta_0 X_0 + \theta_1 X_1 + \theta_2 X_2 + \theta_3 X_1 X_2 + \varepsilon$, and estimating models were $y = \beta_0 X_0 + \beta_1 X_1 + \varepsilon'$.

Generating Model	θ_0	θ_1	θ_2	θ_3	$r(X_1, X_2)$	X_2 spatially structured	Less biased $\beta_1(\tau)$
Additive	1.0	0.41	0.005	0.0000	0.00	No	τ similar
Interference	1.0	0.41	0.0	-0.0001	0.00	No	increasing τ
Facilitation	1.0	0.01	0.0	0.0001	0.00	No	decreasing τ
Interference	1.0	0.41	0.0	-0.0001	0.56	No	increasing τ
Interference	1.0	0.41	0.0	-0.0001	0.92	No	increasing τ
Interference	1.0	0.41	0.0	-0.0001	0.00	Yes	increasing τ

The τ^{th} regression quantile ($0 \leq \tau \leq 1$) of the generating model was defined as $Q_y(\tau | X_0, X_1, X_2, X_1X_2) = \theta_0(\tau)X_0 + \theta_1 X_1 + \theta_2 X_2 + \theta_3 X_1 X_2$, where $\theta_0(\tau) = \theta_0 + F_\varepsilon^{-1}(\tau)$ and F_ε^{-1} was the inverse of the cumulative distribution of the errors. This is just a conventional homoscedastic linear regression model where all parameters other than the intercept (θ_0) are the same for all quantiles τ, i.e. parallel hyperplanes (Cade et al. 1999). The τ^{th} regression quantile of the estimating model where the effect of the unmeasured covariate X_2 was not estimable was $Q_y(\tau | X_0, X_1,) = \beta_0(\tau)X_0 + \beta_1(\tau)X_1$. In the estimating model both the intercept $\beta_0(\tau)$ and slope $\beta_1(\tau)$ for the measured covariate may vary with the quantile τ because the modified error term $\varepsilon' = \varepsilon + \theta_2 X_2 + \theta_3 X_1 X_2$ included the additive random component and a multiplicative component that was a function of the measured covariate X_1, potentially creating mixture distributions that were not identically distributed. Because the parameters of these mixture distributions were not necessarily identifiable, I took a random sample of N = 10,000 and treated this as a large, finite population. I compared regression quantiles for $\beta_0(\tau)$ and $\beta_1(\tau)$ from the estimating model with $\theta_0(\tau)$ and $\theta_1(\tau)$ from the generating model for the finite populations to examine differences in effects associated with the habitat variable (X_1) due to different interaction effects and correlations with the unmeasured variable (X_2).

Sampling distributions of estimates and associated rankscore test statistics (Chapter 2) for the estimating model with heterogeneous mixture distributions were evaluated by taking 1,000 samples of n = 20, 30, 60, 90, 150, and 300 without replacement from the finite populations of N = 10,000. Rankscore tests evaluated were the asymptotic Chi-square distributed T, the permutation F, and the double permutation F for null models constrained through the origin (Chapter 2). The large, finite population can be thought of as 10,000 100-ha blocks occurring on a landscape of 100 × 100 km extent.

Spatial structuring was accomplished by relating the unmeasured limiting factor X_2 to latitude (*LAT*) and longitude (*LONG*) coordinates for the center of 10,000 square blocks on a 100 × 100 grid. I used a cubic polynomial spatial trend surface model (Borcard et al. 1992, Legendre et al. 1997) on mean centered *LAT* and *LONG* coordinates (-50 to 50, μ = 0) with $X_2 = 2{,}000 + 4.5LONG + 7.5LAT + 0.1LONG^2 - 0.2LAT^2 + 0.005LONG^3 + \varepsilon$, with ε uniformly distributed (-900, 900) to yield an $R^2 = 0.426$ with the least squares regression estimate of the spatial trend surface (Fig.4.1). The spatially structured X_2 had values ranging from 0 to 4,000 and was uncorrelated with X_1. Obviously, similar spatial structuring could have been induced in either the response variable y or in the measured habitat variable X_1.

3. Patterns of Effects Due to Confounding with Unmeasured Processes

To explore patterns associated with missing information on some important generating process, I started with the simplest case of no spatial structuring, no correlation between the measured (X_1) and unmeasured (X_2) variables, and no interaction effect ($\theta_3 = 0.0$) in the additive generating model $y = \theta_0X_0 + \theta_1X_1 + \theta_2X_2 + \theta_3X_1X_2 + \varepsilon$; $\theta_0 = 1.0$, $\theta_1 = 0.41$, $\theta_2 = 0.005$, and $\theta_3 = 0.0$. If the estimating model $y = \beta_0X_0 + \beta_1X_1 + \varepsilon'$ was used because X_2 was unmeasured, all the unexplained variation due to θ_2 was additive in the new error term $\varepsilon' = \varepsilon + 0.005\times X_2$, which caused differences between the intercept parameters for the generating, θ_0, and estimating, β_0, models but only small differences between the slope parameters θ_1 and β_1 (Fig.4.2). The estimating model had homogeneous variances like the generating model but with bias in intercepts and no bias in slopes. Thus, rates of change due to X_1 estimated for any quantile or for the mean with least squares regression would be similar in repeated random sampling. This example clarified why the heterogeneous constraint patterns investigated by Terrell et al. (1996), Cade et al. (1999), and Huston (2002) imply that there must be more than just additive effects between the measured and unmeasured processes in the

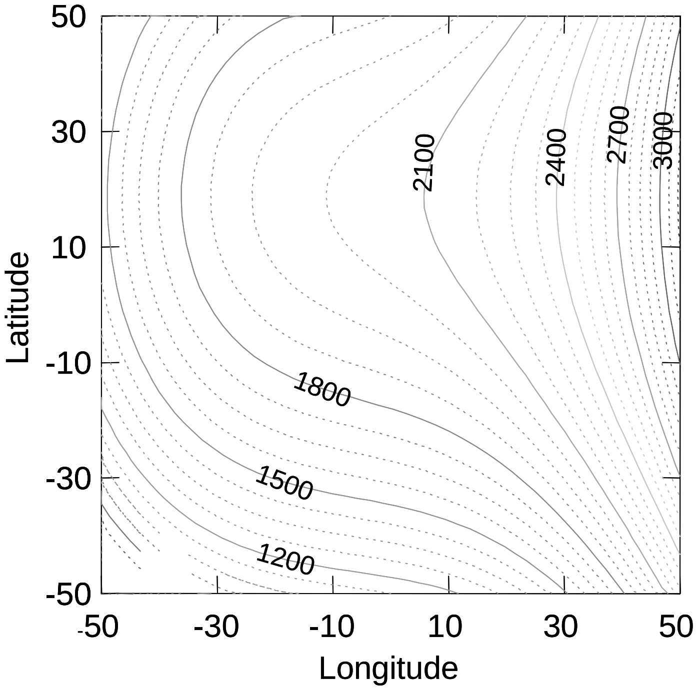


Figure 4.1. Cubic polynomial spatial trend surface used in simulations to generate the values of X_2, an unmeasured nonhabitat variable that was not estimable in simulations. Surface plotted is for the expected value of $X_2 = 2{,}000 + 4.5LONG + 7.5LAT + 0.1LONG^2 - 0.2LAT^2 + 0.005LONG^3 + \varepsilon$, with ε uniformly distributed (-900, 900).

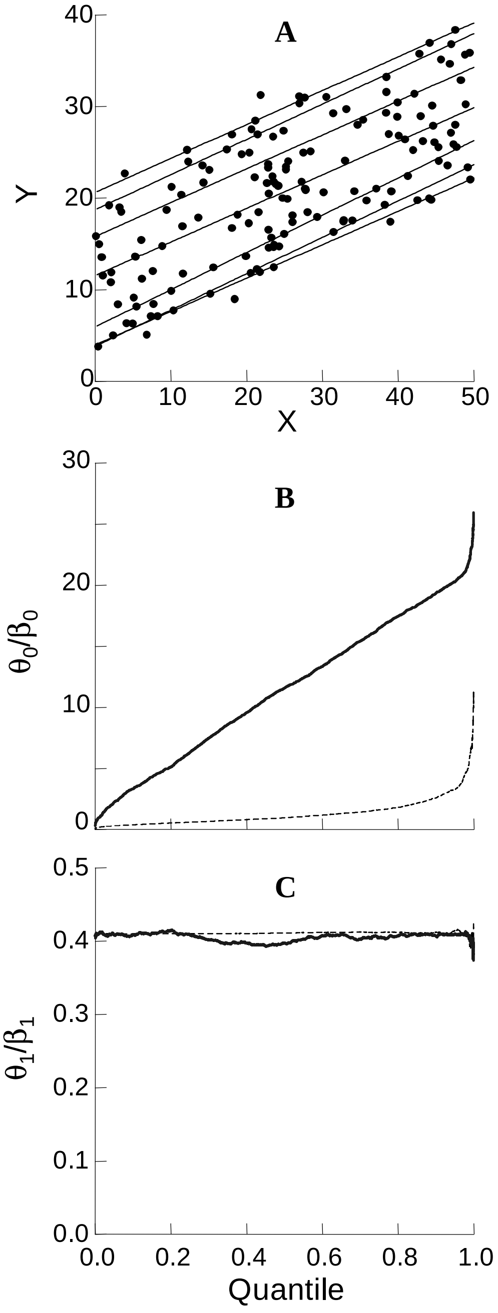


Figure 4.2. (A) A sample (n = 150) from the N = 10,000 population for the additive generating model $y = \theta_0 X_0 + \theta_1 X_1 + \theta_2 X_2 + \theta_3 X_1 X_2 + \varepsilon$; $\theta_0 = 1.0$, $\theta_1 = 0.41$, $\theta_2 = 0.005$, $\theta_3 = 0.0$, and ε lognormally distributed (median = 0, σ = 0.75). Lines plotted are for regression quantile estimates ($\tau \in \{0.95, 0.90, 0.75, 0.50, 0.25, 0.10, 0.05\}$) when the estimating model is $y = \beta_0 X_0 + \beta_1 X_1 + \varepsilon'$. (B) Shows $\beta_0(\tau)$ and $\theta_0(\tau)$ deviating more at higher quantiles (τ) (θ's are dashed and β's are solid lines). The ordinary Least squares β_0 = 11.326. (C) Shows $\beta_1(\tau)$ and $\theta_1(\tau)$ deviating slightly for any quantile. The ordinary least squares β_1 = 0.381.

linear model.

A multiplicative interference interaction ($\theta_3 < 0.0$) model with no spatial structuring and no correlation between measured and unmeasured variables with $\theta_0 = 1.0$, $\theta_1 = 0.41$, $\theta_2 = 0.0$, and $\theta_3 = -0.0001$ was estimated without information about the interaction effect, yielding an increasing variance pattern similar to those discussed by Terrell et al. (1996), Thomson et al. (1996), Cade et al. (1999), and Huston (2002). Here both intercept (β_0) and slope (β_1) parameters for the different quantiles of the estimating model were biased relative to θ_0 and θ_1 of the generating model (Fig.4.3). Values of $\beta_1(\tau)$ at higher quantiles ($\tau > 0.90$) were the least biased relative to θ_1. By algebraically reexpressing the interaction effects associated with the measured habitat variable in the generating model as $(\theta_1 + \theta_3 X_2)X_1$ it is possible to explain the source of differing values of $\beta_1(\tau)$ across quantiles. The $\beta_1(\tau)$ for higher quantiles are effects of the measured habitat variable, θ_1 minus a small quantity, when the unmeasured variable X_2 is close to its minimum of zero ($-0.0001 \times X_2 \rightarrow 0$). The $\beta_1(\tau)$ for lower quantiles are effects of the measured habitat variable, θ_1 minus a large quantity, when the unmeasured variable X_2 is close to its maximum of 4,000 ($-0.0001 \times X_2 \rightarrow -0.40$). The lognormal error distribution used in this example resulted in a mixture distribution ($\varepsilon' = \varepsilon + -0.0001 X_1 X_2$) that prevented the convergence of $\beta_1(\tau)$ with θ_1 at highest quantiles. However, when this example was simulated with the uniform error distribution, $\beta_1(\tau)$ converged with θ_1 at the highest quantiles. The lesson is that we can never be sure of the magnitude of bias associated with effects estimated for some upper quantiles since in applications we will never know the exact distributional form of the errors.

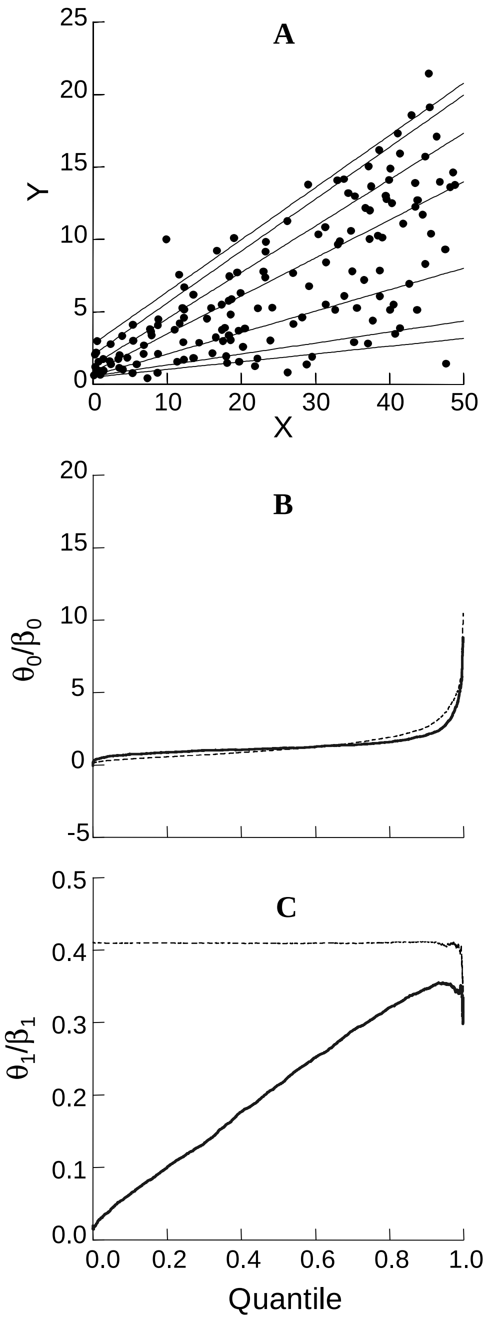


Figure 4.3. (A) A sample (n = 150) from the N = 10,000 population for the interference interaction generating model $y = \theta_0 X_0 + \theta_1 X_1 + \theta_2 X_2 + \theta_3 X_1 X_2 + \varepsilon$; $\theta_0 = 1.0$, $\theta_1 = 0.41$, $\theta_2 = 0.0$, $\theta_3 = -0.0001$, and ε lognormally distributed (median = 0, $\sigma = 0.75$). Lines plotted are for regression quantile estimates ($\tau \in \{0.95, 0.90, 0.75, 0.50, 0.25, 0.10, 0.05\}$) when the estimating model is $y = \beta_0 X_0 + \beta_1 X_1 + \varepsilon'$. (B) Shows $\beta_0(\tau)$ and $\theta_0(\tau)$ deviating slightly across the quantiles (τ) (θ's are dashed and β's are solid lines). The ordinary least squares $\beta_0 = 1.438$. (C) Shows $\beta_1(\tau)$ and $\theta_1(\tau)$ deviating less for higher quantiles. The ordinary least squares $\beta_1 = 0.204$.

However, we can be confident that estimates for upper quantiles are less biased than those for lower quantiles or for the mean (least squares regression) when the assumption about interference interactions is reasonable, because higher quantile estimates of $\beta_1(\tau)$ include interaction effects at lower values of the unmeasured variable.

A multiplicative facilitation interaction ($\theta_3 > 0.0$) model with no spatial structuring and no correlation between measured and unmeasured processes with $\theta_0 = 1.0$, $\theta_1 = 0.01$, $\theta_2 = 0.0$, and $\theta_3 = 0.0001$ was generated and then estimated without information about the interaction effect. This yielded an increasing variance pattern similar to the previous example for the interference interaction (Fig. 4.4). Here, $\beta_1(\tau)$ at lower quantiles ($\tau < 0.05$) were the least biased relative to θ_1. Algebraically reexpressing the interaction effects associated with the measured habitat variable as in the previous example, indicates that $\beta_1(\tau)$for higher quantiles are effects of the measured habitat variable, θ_1 plus a large quantity, when the unmeasured variable X_2 is close to its maximum of 4,000 ($0.0001 \times X_2 \rightarrow 0.40$). The $\beta_1(\tau)$ for lower quantiles are effects of the measured habitat variable, θ_1 plus a small quantity, when the unmeasured variable X_2 is close to its minimum of zero ($0.0001 \times X_2 \rightarrow 0$). The lesson is that selecting lower or upper quantiles to provide less biased estimates of effects of some measured habitat variable when effects of unmeasured variables are negligible (i.e, habitat as the limiting constraint) is critically dependent on the assumed type of interaction (+ for facilitation or - for interference) between the variables. While interference interactions may be more common in ecological systems, facilitation

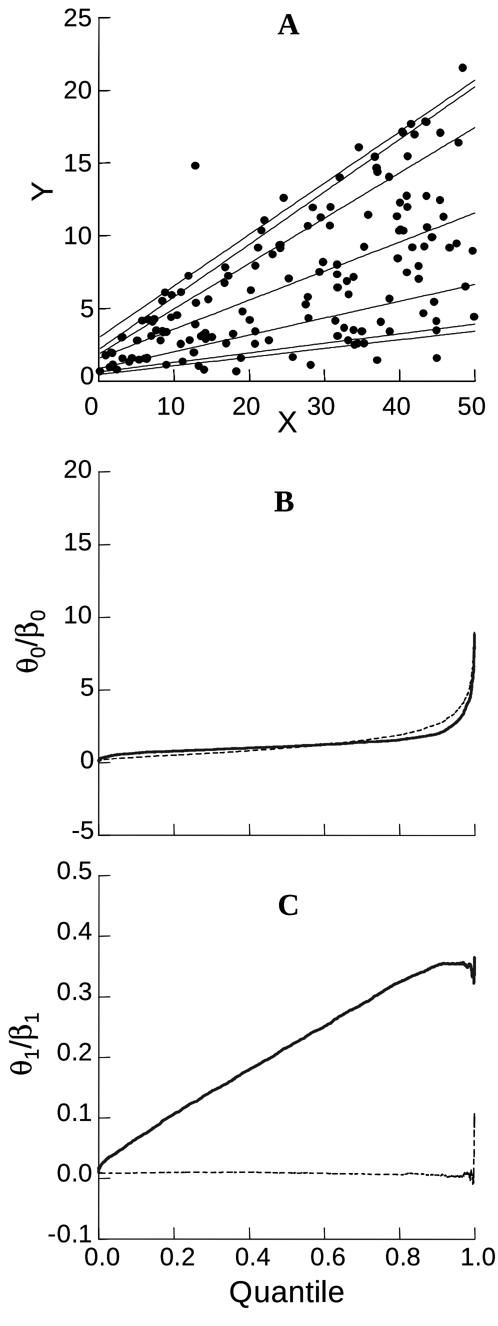


Figure 4.4. (A) A sample (n = 150) from the N = 10,000 population for the facilitation interaction generating model $y = \theta_0 X_0 + \theta_1 X_1 + \theta_2 X_2 + \theta_3 X_1 X_2 + \varepsilon$; $\theta_0 = 1.0$, $\theta_1 = 0.01$, $\theta_2 = 0.0$, $\theta_3 = 0.0001$, and ε lognormally distributed (median = 0, $\sigma = 0.75$). Lines plotted are for regression quantile estimates ($\tau \in \{0.95, 0.90, 0.75, 0.50, 0.25, 0.10, 0.05\}$) when the estimating model is $y = \beta_0 X_0 + \beta_1 X_1 + \varepsilon'$. (B) Shows $\beta_0(\tau)$ and $\theta_0(\tau)$ deviating slightly across the quantiles (τ) (θ's are dashed and β's are solid lines). The ordinary least squares $\beta_0 = 1.307$. (C) Shows $\beta_1(\tau)$ and $\theta_1(\tau)$ deviating less for lower quantiles. The ordinary least squares $\beta_1 = 0.211$.

interactions might occur in situations where over short time spans the resources provided by the habitat were insufficient to support the population, e.g., salmonid populations reproducing in streams that lack the food or cover resources to sustain them throughout their life cycle. A determination of whether interference or facilitation interactions are likely requires knowledge obtained from sources other than the data being analyzed.

A slightly more complicated interference interaction model was simulated with varying amounts of correlation (r = 0.56 and 0.92) between the measured habitat variable X_1 and the unmeasured variable X_2 with $\theta_0 = 1.0$, $\theta_1 = 0.41$, $\theta_2 = 0.0$, and $\theta_3 = -0.0001$. The estimating model without the interaction effect now yields an increasing variance pattern with slight nonlinearity evident for r = 0.56 and a more homogeneous variance pattern with stronger nonlinearity for r = 0.92 (Fig. 4.5). The source of the nonlinearity is explained by recognizing that the interaction effect is $X_1 \times X_2$ but because of the correlation structure X_2 is a function of X_1, e.g., r = 0.56 was achieved by the function $X_2 = 1{,}200 + 32.0 \times X_1$ + uniform random number [-1200, 1200]. So the interaction effect was a function involving X_1^2, a quadratic polynomial. Depending on the sign of the interaction coefficient (θ_3) and sign of the correlation (r), nonlinear functions may curve upward (+, + and -,-) or downward (+, - and -, +). The lesson is that correlation between measured habitat resources and unmeasured variables results in nonlinear response relationships, the stronger the correlation the greater the nonlinearity and less heterogenous the response. This suggests that some surrogate variable that is strongly correlated with the unmeasured variables might help account

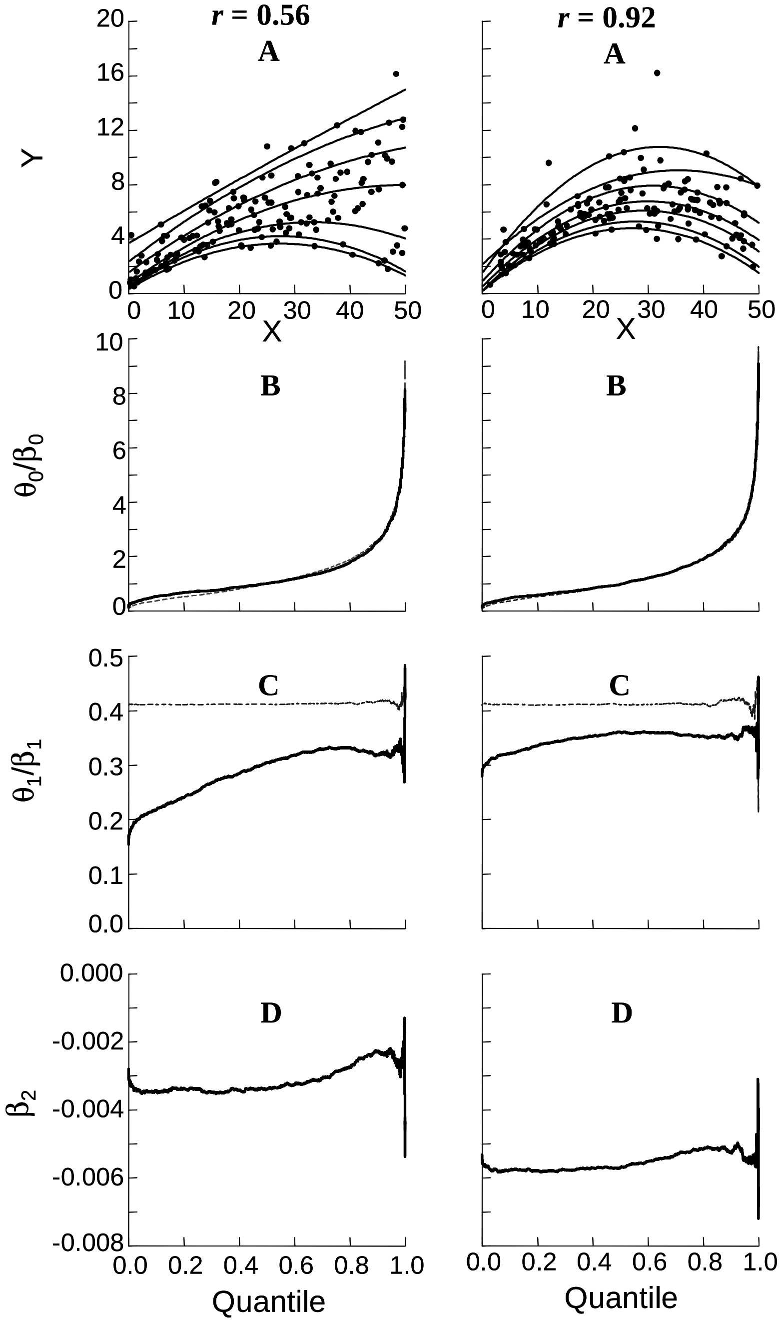


Figure 4.5. (A) A sample (n = 150) from the N = 10,000 population of grid cells from the interference interaction generating model as in Figure 4.3 but with $r(X_1, X_2)$ = 0.56 and 0.92. Lines plotted are for selected regression quantile estimates when the estimating model is $y = \beta_0 X_0 + \beta_1 X_1 + \beta_2 X_1^2 + \varepsilon'$ because X_2 was not measured. (B) Shows $\beta_0(\tau)$ and $\theta_0(\tau)$ deviating slightly for some quantiles (τ) (θ's have dashed and β's have solid lines). (C) Shows $\beta_1(\tau)$ and $\theta_1(\tau)$ deviating less for higher quantiles and for $r(X_1, X_2)$ = 0.92. (D) Shows $\beta_2(\tau)$ across quantiles with more negative estimates for $r(X_1, X_2)$ = 0.92 indicating greater nonlinearity.

for some of the variation in the modeled relationships.

The spatial coordinates of sample locations are a potential set of surrogate variables for unmeasured processes that are spatially structured. An interference interaction model was simulated with no correlation between measured and unmeasured variables but with the unmeasured variable related to latitudinal and longitudinal coordinates (Fig. 4.1) and $\theta_0 = 1.0$, $\theta_1 = 0.41$, $\theta_2 = 0.0$, and $\theta_3 = -0.0001$. The estimating model $y = \beta_0 X_0 + \beta_1 X_1 + \beta_2 X_1 \times LAT + \beta_3 X_1 \times LONG + \beta_4 X_1 \times LAT^{2} + \beta_5 X_1 \times LONG^{2} + \beta_6 X_1 \times LONG^{3} + \varepsilon'$ had relatively homogeneous effects across quantiles for the interaction of the measured habitat variable with cubic polynomial terms (β_2 - β_6) for the spatial trend surface (Fig. 4.6). Variation in $\beta_1(\tau)$ across quantiles was evident for the measured habitat variable with less bias relative to θ_1 at higher quantiles. Notice by comparing $\beta_1(\tau)$ in Figure 4.6, where some of the effect of the unmeasured variable was accounted for by the spatial trend, with $\beta_1(\tau)$ in Figure 4.3, where it was not, that variation in rate parameters across quantiles was less for the former model although bias was greater for parameters at higher quantiles. Stronger spatial structuring of the unmeasured variable (X_2) would have produced less variation in $\beta_1(\tau)$ across quantiles and less bias relative to θ_1. However, the amount of variance explained ($R^2 = 0.426$) with the spatial trend surface simulated was typical of the better results achieved in ecological investigations (e.g., Legendre et al. 1997).

In applications of the spatial trend surface, interactions of environmental covariates and cubic polynomial terms usually were not estimated (Borcard et al. 1992, Legendre et al. 1997, Legendre and Legendre 1998, Lichstein et al. 2002). My

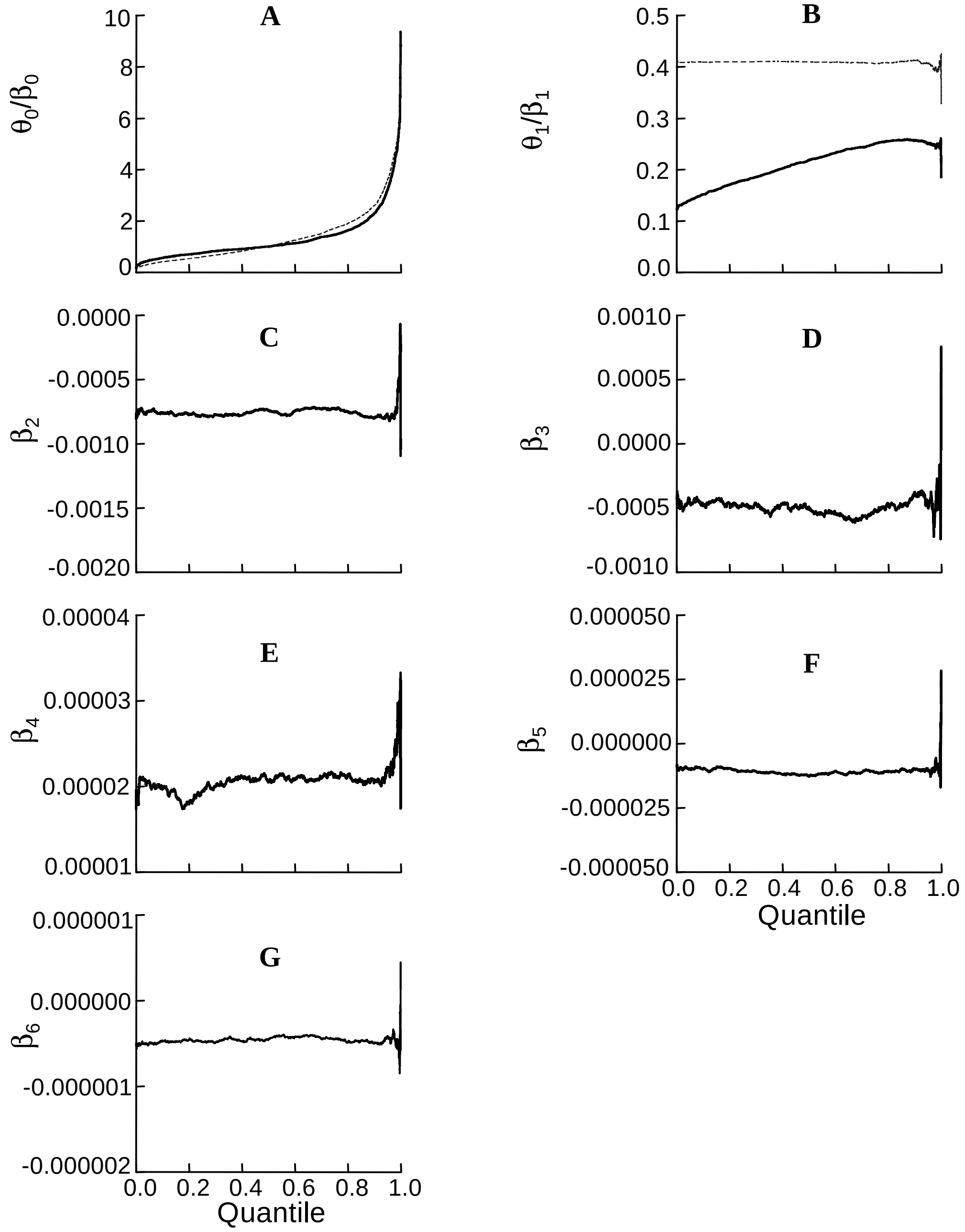


Figure 4.6. Parameters for the N = 10,000 population of grid cells from the interference interaction generating model as in Figure 4.3 but with $X_2 = 2{,}000 + 4.5LONG + 7.5LAT + 0.1LONG^2 - 0.2LAT^2 + 0.005LONG^3 + \varepsilon$, with ε uniformly distributed (-900, 900) (θ's have dashed lines); and for the estimating model $y = \beta_0 X_0 + \beta_1 X_1 + \beta_2 X_1 \times LONG + \beta_3 X_1 \times LAT + \beta_4 X_1 \times LONG^2 + \beta_5 X_1 \times LAT^2 + \beta_6 X_1 \times LONG^3 + \varepsilon'$ (β's have solid lines) used because X_2 was not measured. (A) Shows $\theta_0(\tau)$ and $\beta_0(\tau)$ deviating slightly for some quantiles (τ). (B) Shows $\theta_1(\tau)$ and $\beta_1(\tau)$ deviating less for higher quantiles. (C) - (G) show relatively homogeneous effects of $\beta_2(\tau)$, $\beta_3(\tau)$, $\beta_4(\tau)$, $\beta_5(\tau)$, and $\beta_6(\tau)$ across quantiles for the interactions with the cubic polynomial spatial trend.

simulations suggest that estimating these interactions might be reasonable when using spatial structure as a surrogate for important unmeasured processes. Here I also simulated an estimating model with a full cubic polynomial without interactions with the measured habitat variable, $y = \beta_0 X_0 + \beta_1 X_1 + \beta_2 LAT + \beta_3 LONG + \beta_4 LAT^2 + \beta_5 LONG^2 + \beta_6 LAT \times LONG + \beta_7 LAT^2 \times LONG + \beta_8 LAT \times LONG^2 + \beta_9 LAT^3 + \beta_{10} LONG^3 + \varepsilon'$. Some of the parameters for the polynomial trend surface terms exhibited heterogeneity across quantiles, but $\beta_1(\tau)$ for effects of the measured habitat variable had similar differences across quantiles as in the more appropriate trend surface with interaction effects model. The lesson is that with reasonable amounts of spatial structuring of unmeasured processes it is still likely that considerable heterogeneity in responses remains in models that incorporate a spatial trend surface. Bias in parameters is not necessarily lessened by including spatial trend.

4. Performance of Regression Quantile Tests for Models with Hidden Bias

Confidence intervals for regression quantile estimates commonly are computed by inverting rankscore testing procedures. Interval coverage for estimates made with hidden bias in the models was estimated with a simulation experiment evaluating Type I error rates of the asymptotic Chi-square T and permutation F rankscore tests (Chapter 2). One thousand random samples for n = 20, 30, 60, 90, 150, and 300 were drawn from the finite population of N = 10,000 blocks for the interference interaction generating model with no spatial structure, no correlation between measured (X_1) and unmeasured (X_2) variables, and $\theta_0 = 1.0$, $\theta_1 = 0.41$, $\theta_2 = 0.0$, and $\theta_3 = -0.0001$ as in Figure 4.3. Estimates of $\beta_0(\tau)$ and $\beta_1(\tau)$ were made for each sample, and null hypotheses H_0: $\beta_0(\tau) = \xi_0(\tau)$, and H_0: $\beta_1(\tau) = \xi_1(\tau)$ were evaluated, where $\xi_0(\tau)$ were the parameter values 5.0662, 2.7230, 2.0720, 1.4503, 1.1368, 0.9186, 0.7304, 0.5935, and 0.3784 and $\xi_1(\tau)$ were the parameter values 0.3445, 0.3533, 0.3464, 0.3034, 0.2139, 0.1170, 0.0628, 0.0431, and 0.0227 corresponding to τ = 0.99, 0.95, 0.90, 0.75, 0.50, 0.25, 0.10, 0.05, and 0.01 quantiles, respectively (Fig. 4.3). This simulation approach evaluated whether the confidence interval coverage estimated by inverting the rankscore tests included the parameter values $\beta_0(\tau)$ and $\beta_1(\tau)$ with the stated confidence level (1 - α) given that the error distribution included effects of the unmeasured variable. Section 3 already established the degree that $\beta_0(\tau)$ and $\beta_1(\tau)$ were biased relative to $\theta_0(\tau)$ and $\theta_1(\tau)$.

Unweighted estimates and rankscore tests provided liberal error rates for H_0: $\beta_1(\tau) = \xi_1(\tau)$ for $\tau < 0.90$, consistent with simulations when the model form was completely specified and heterogeneity was >5 standard deviations across the domain of X (Chapter 2). It was only at higher quantiles $\tau \geq 0.90$, where there was a reduced rate of change between $\beta_1(\tau)$ (see Fig. 4.3), that unweighted estimates and rankscore tests provided reasonable coverage (Fig. 4.7). The permutation F rankscore test maintained better Type I error rates than the T rankscore test for smaller samples at more extreme quantiles (Fig. 4.7), similar to simulations without hidden bias (Chapter 2). Unweighted estimates and T rankscore tests provided good Type I error rates for H_0: $\beta_0(\tau) = \xi_0(\tau)$ across all but the most extreme quantiles (τ = 0.01 and 0.99), whereas F rankscore tests had slightly liberal error rates because the permutation structure used did not account for all the sampling variability when null models were forced through the origin (Fig.4.8), similar to simulations in Chapter 2. The double permutation scheme improved Type I error rates for the permutation *F* test as demonstrated below for weighted estimates of $\beta_1(\tau)$.

Weighted estimates and rankscore tests were simulated by constructing weights based on the regression quantile parameters for the N = 10,000 finite population (Fig. 4.3). The pattern of increasing $\beta_1(\tau)$ with increasing τ was not a simple location-scale form because differences in $\beta_1(\tau)$ were not constant across all τ, although they differed by a fairly constant amount for $0.20 \leq \tau \leq 0.80$ (Fig. 4.3). A variant of the bandwidth approach based on changes in $\beta_0(\tau)$ and $\beta_1(\tau)$ near the quantile (τ) of interest (Koenker and Machado 1999) was used to provide weights for weighted estimates and rankscore tests in simulations. Weights were, thus, based on the N = 10,000 population and not estimated for different samples to avoid undue complexity in the simulation experiment. Weights were computed by taking the average pairwise difference between $\beta_0(\tau)$ and between $\beta_1(\tau)$ in an interval of $\tau \pm 0.01$ for τ = 0.05, 0.10, 0.90, and 0.95 and in an interval $\tau \pm 0.005$ for τ = 0.01 and 0.99. For τ = 0.25, 0.50, and 0.75 there was almost constant rate of change in the parameters, and weights were computed based on pairwise differences in the interval τ = [0.25, 0.75]. Weights, $w(\tau)$, were the reciprocal of the average pairwise differences divided by the associated interval width used in their computation (0.01, 0.02, or 0.50): $w(0.99) = (48.825 + 0.377X_1)^{-1}$, $w(0.95) = (9.195 - 0.041X_1)^{-1}$, $w(0.90) = (3.270 + 0.051X_1)^{-1}$, $w(0.75) = w(0.50) = w(0.25) = (0.286 +.0.110X_1)^{-1}$, $w(0.10) = (0.900 + 0.129X_1)^{-1}$, $w(0.05) = (0.774 + 0.169X_1)^{-1}$, and $w(0.01) = (2.589 + 0.288X_1)^{-1}$. Weights, $w(\tau)$, were then multiplied by y and X to

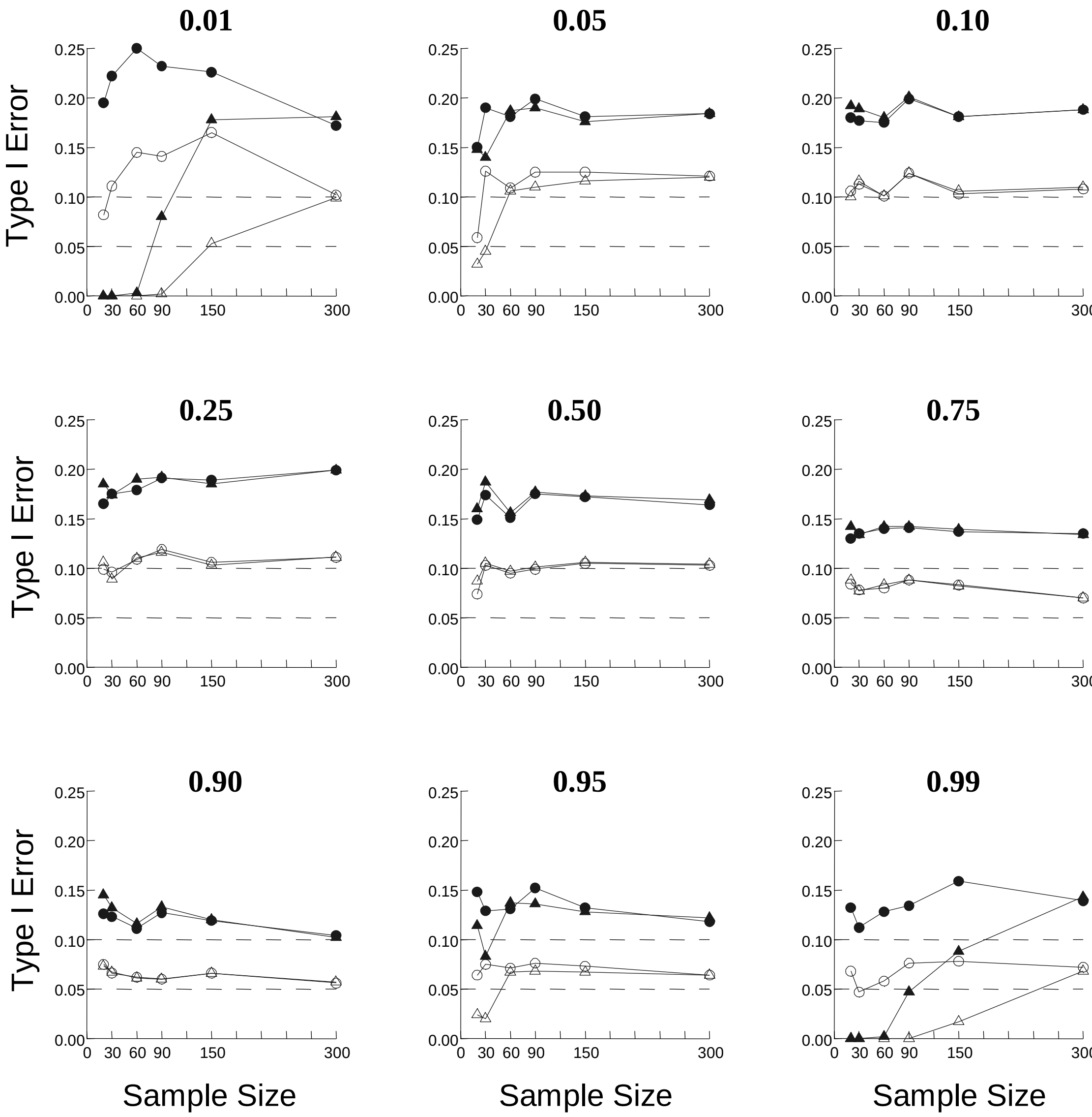


Figure 4.7. Estimated Type I error rates for $\alpha = 0.05$ (open) and 0.10 (solid); for the T (triangles) and permutation F (circles) rankscore tests for H_0: $\beta_1(\tau) = \xi_1(\tau)$ in the estimating model $y = \beta_0 X_0 + \beta_1 X_1 + \varepsilon'$, where $\xi_1(\tau)$ were the parameter values 0.344, 0.353, 0.346, 0.303, 0.214, 0.117, 0.063, 0.043, and 0.023 for τ = 0.99, 0.95, 0.90, 0.75, 0.50, 0.25, 010, 0.05, and 0.01 for the N = 10,000 grid cells generated by the model in Figure 4.3; and for n = 20, 30, 60, 90, 150, and 300. 1,000 random samples were used for each n.

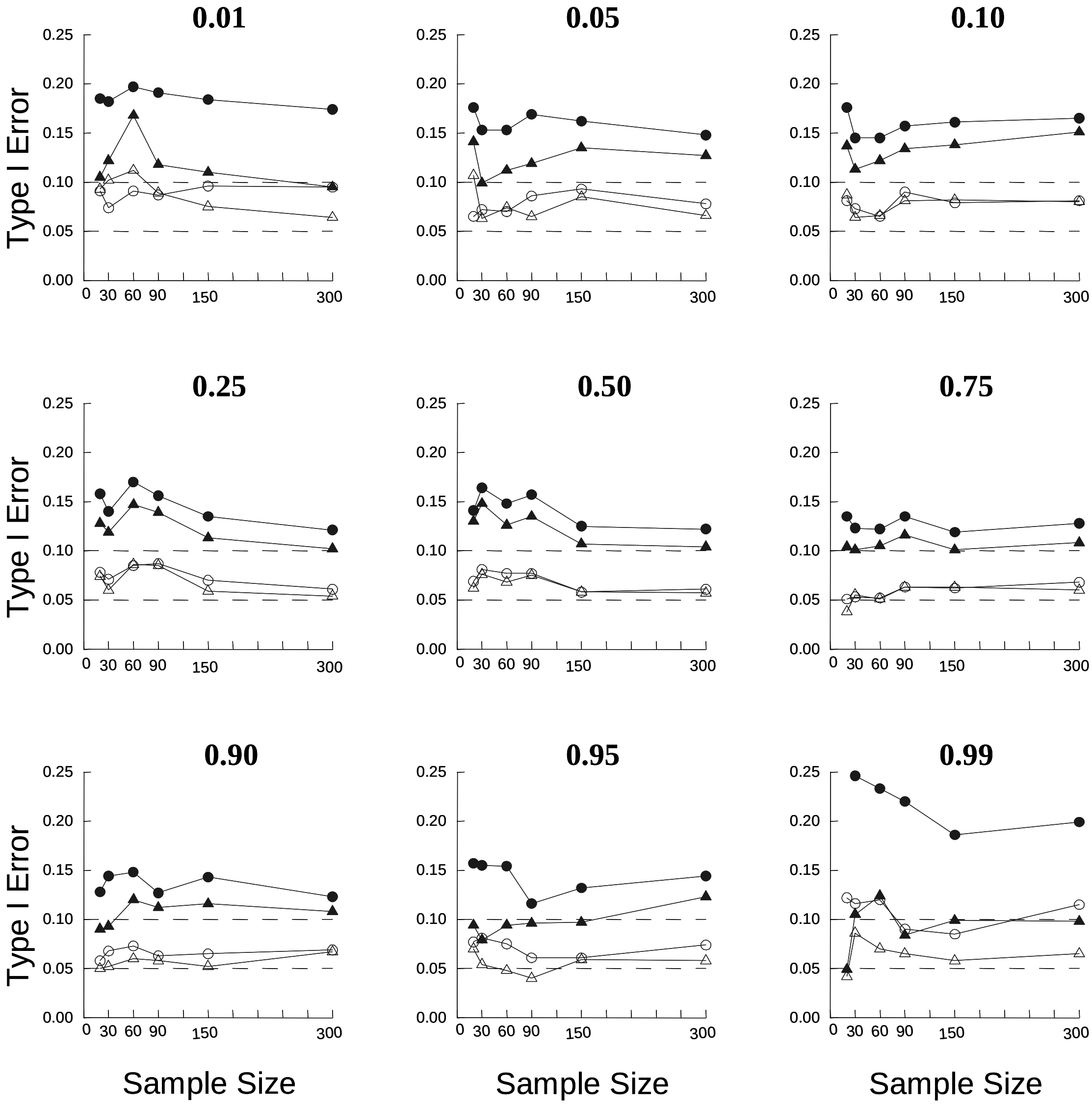


Figure 4.8. Estimated Type I error rates for $\alpha = 0.05$ (open) and 0.10 (solid); for the *T* (triangles) and permutation *F* (circles) rankscore tests for H_0: $\beta_0(\tau) = \xi_0(\tau)$ in the estimating model $y = \beta_0 X_0 + \beta_1 X_1 + \varepsilon'$, where $\xi_0(\tau)$ were the parameter values 5.066, 2.723, 2.072, 1.450, 1.137, 0.919, 0.593, and 0.378 for τ = 0.99, 0.95, 0.90, 0.75, 0.50, 0.25, 010, 0.05, and 0.01 for the N = 10,000 grid cells generated by the model in Figure 4.3; and for n = 20, 30, 60, 90, 150, and 300. 1,000 random samples were used for each n.

compute weighted regression quantile estimates and their associated rankscore tests as in Chapter 2.

Type I error rates for H_0: $\beta_1(\tau) = \xi_1(\tau)$ were maintained for the weighted *T* test across all quantiles except for $\tau = 0.01$ (Fig. 4.9). The weighted *F* test had slightly conservative error rates compared to the weighted *T* test except for $\tau = 0.90$ where they both maintained correct levels because rates of change in adjacent $\beta_1(\tau)$ and affects of the weighting were minimal. At extreme quantiles and smaller *n* the weighted *T* test became extremely conservative compared to the weighted *F* test. The permutation scheme based on the weighted regression quantile estimates clearly would benefit from some adjustment because weighted estimates of the null model were forced through the origin. The double permutation scheme (Chapter 2) applied to the *F* test provided improved Type I error rates compared to the standard permutation scheme, although probabilities became conservative at more extreme τ and smaller *n* (Fig. 4.10). Weighting provided minor improvements to error rates for H_0: $\beta_0(\tau) = \xi_0(\tau)$ for the *T* test and little improvement for the standard permutation *F* test for most quantiles (Fig. 4.11). Weighting actually made the error rates for the 0.01 quantile more unstable. The double permutation scheme provided minor improvements for the *F* test for this hypothesis.

Type I error rates for the cubic polynomial trend surface were evaluated for the interference interaction model with no spatial structuring (Fig. 4.3). Estimates of $\beta_0(\tau)$, $\beta_1(\tau)$, $\beta_2(\tau)$, $\beta_3(\tau)$, $\beta_4(\tau)$, $\beta_5(\tau)$, $\beta_6(\tau)$, $\beta_7(\tau)$, $\beta_8(\tau)$, $\beta_9\tau)$, and $\beta_{10}(\tau)$, were made for each sample and the null hypothesis H_0: $\beta_2(\tau) = \beta_3(\tau) = ... = \beta_{10}(\tau) = 0$ was tested, where $\beta_2(\tau)$

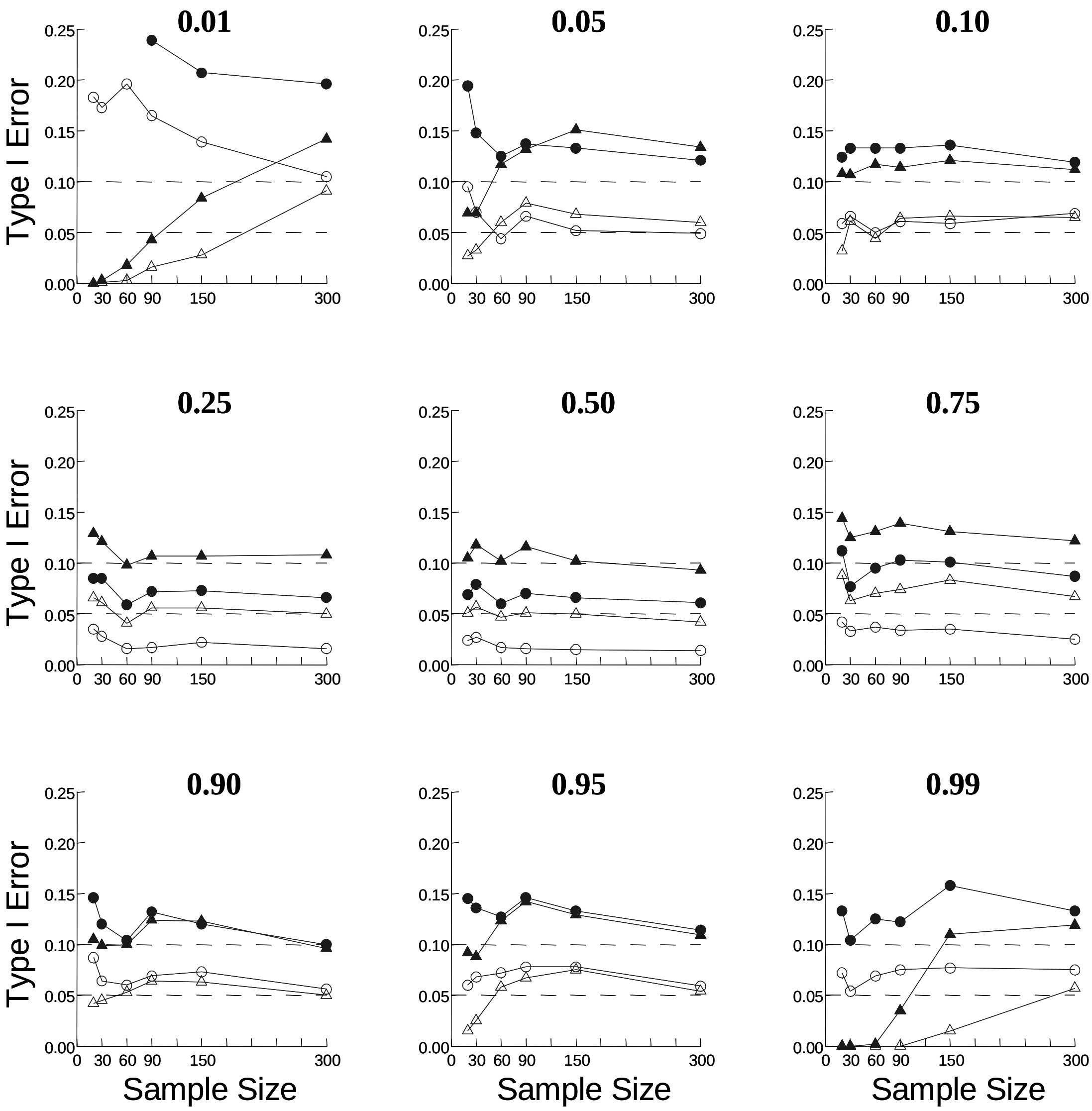


Figure 4.9. Estimated Type I error rates for $\alpha = 0.05$ (open) and 0.10 (solid); for the T (triangles) and permutation F (circles) rankscore tests for H_0: $w\beta_1(\tau) = \xi_1(\tau)$ in the estimating model $wy = w\beta_0 X_0 + w\beta_1 X_1 + w\varepsilon'$, where $\xi_1(\tau)$ were the parameter values 0.341, 0.354, 0.345, 0.302, 0.217, 0.126, 0.068, 0.048, and 0.025 for $\tau = 0.99$, 0.95, 0.90, 0.75, 0.50, 0.25, 010, 0.05, and 0.01 for the $N = 10{,}000$ grid cells generated by the model in Figure 4.3; and for $n = 20$, 30, 60, 90, 150, and 300. 1,000 random samples were used for each n.

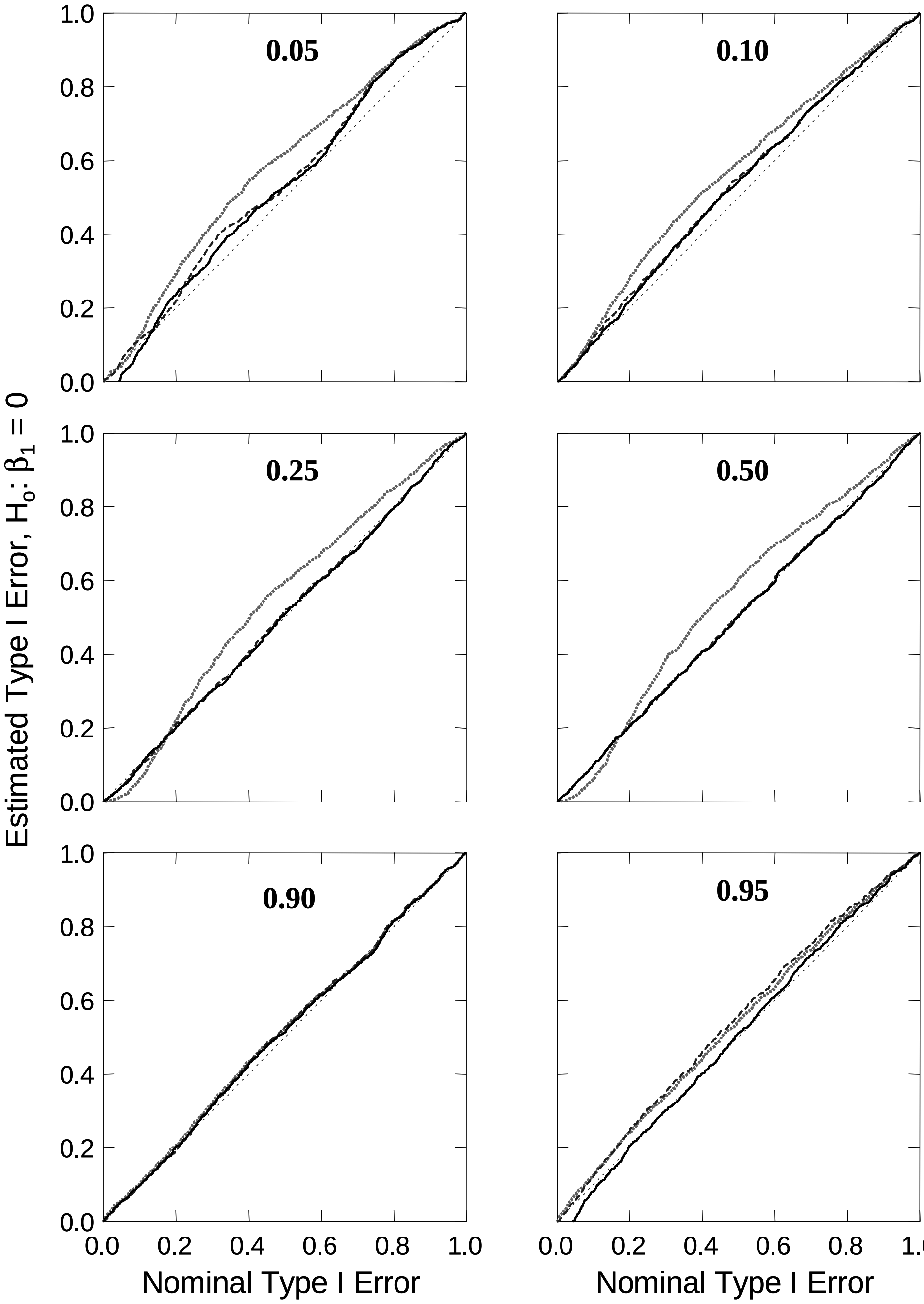


Figure 4.10. Cumulative distributions of 1,000 estimated errors for the Chi-squared distributed *T* (dashed line), permutation *F* (square dotted line), and double permutation *F* (solid line) rankscore tests for H_0: $w\beta_1(\tau) = \xi_1(\tau)$ for $n = 60$ in the estimating model $wy = w\beta_0 X_0 + w\beta_1 X_1 + w\varepsilon'$, where $\xi_1(\tau)$ were the parameter values 0.354, 0.345, 0.217, 0.126, 0.068, and 0.048 for τ = 0.95, 0.90, 0.50, 0.25, 010, and 0.05 for the N = 10,000 grid cells generated by the model in Figure 4.3. Fine dotted line is the expected cdf of a uniform distribution. See text for description of weights $w(\tau)$.

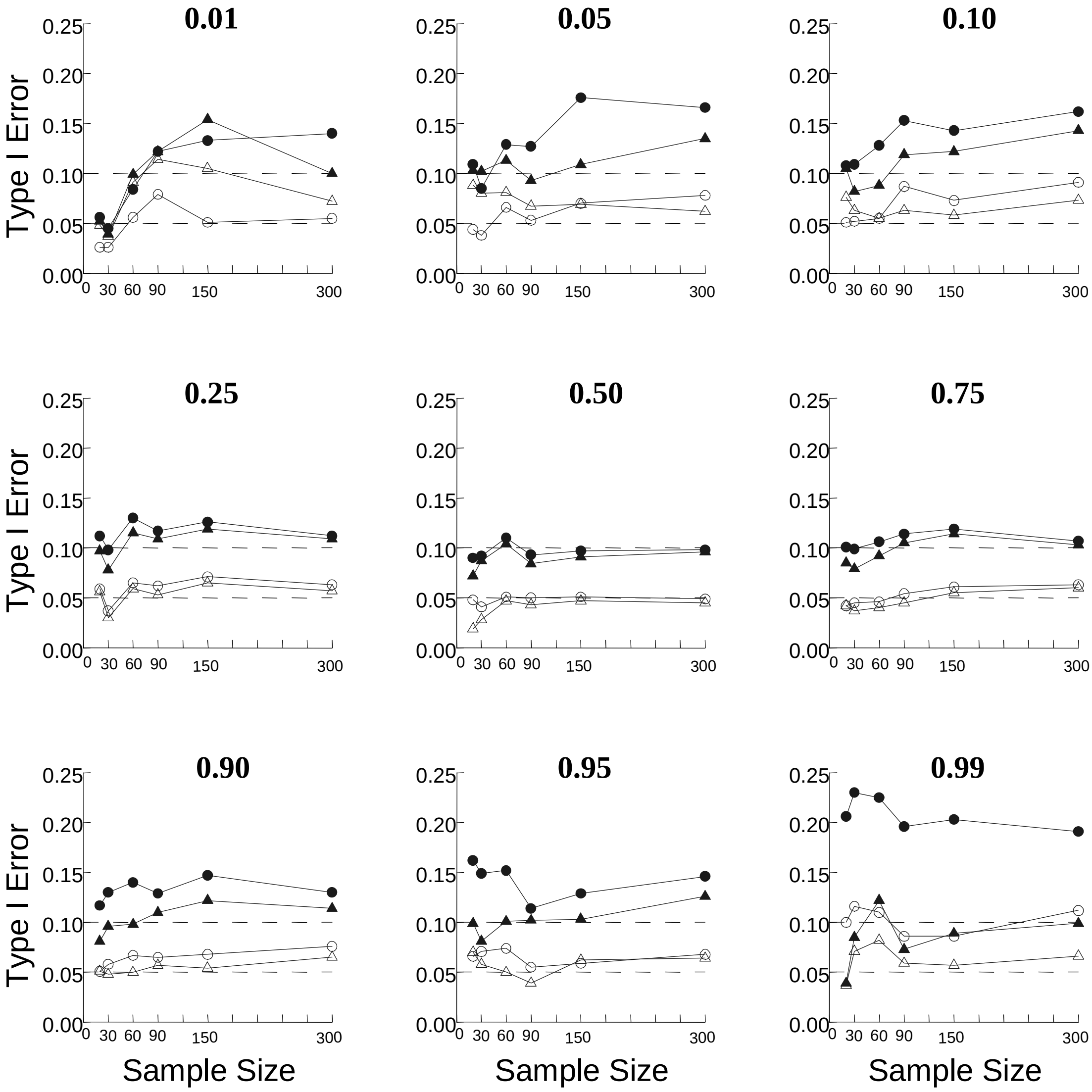


Figure 4.11. Estimated Type I error rates for $\alpha = 0.05$ (open) and 0.10 (solid); for the T (triangles) and permutation F (circles) rankscore tests for H_0: $w\beta_0(\tau) = \xi_0(\tau)$ in the estimating model $wy = w\beta_0 X_0 + w\beta_1 X_1 + w\varepsilon'$, where $\xi_0(\tau)$ were the parameter values 5.134, 2.693, 2.110, 1.478, 1.084, 0.802, 0.645, and 0.355 for τ = 0.99, 0.95, 0.90, 0.75, 0.50, 0.25, 010, 0.05, and 0.01 for the N = 10,000 grid cells generated by the model in Figure 4.3; and for n = 20, 30, 60, 90, 150, and 300. 1,000 random samples were used for each n.

- $\beta_{10}(\tau)$ were parameters (all zero) for the 9 terms of the full cubic polynomial trend surface. Here Type I error rates were well maintained by both the *T* and *F* rankscore tests because under the alternative model there was no relation between the spatial trend surface and the response for any quantile (Fig. 4.12). The permutation evaluation of the *F* statistic provided better Type I error rates than the asymptotic Chi-square evaluation of the *T* statistic for smaller *n* at more extreme quantiles, as also observed for models without hidden bias (Chapter 2).

A small simulation experiment was conducted to evaluate power of the regression quantile estimates and rankscore tests to detect spatial trend surfaces. Samples (n = 1,000) were taken from the spatially structured, interference interaction population of N = 10,000 blocks (Fig. 4.6), and the model $y = \beta_0(\tau)X_0 + \beta_1(\tau)X_1 + \beta_2(\tau)X_1 \times LAT + \beta_3(\tau)X_1 \times LONG + \beta_4(\tau)X_1 \times LAT^2 + \beta_5(\tau)X_1 \times LONG^2 + \beta_6(\tau)X_1 \times LONG^3$ was estimated and rankscore tests conducted for H_0: $\beta_2(\tau) = \beta_3(\tau) = \ldots = \beta_6(\tau) = 0$. Because the simulated effect of the spatial trend surface was homogeneous across quantiles, no weighting was used in the simulations. Power greater than 80% with $\alpha = 0.05$ was achieved for $n \geq 150$ for τ = 0.05 - 0.90. Power was 52% for τ = 0.95, 30% for τ = 0.01, and 7% for τ = 0.99 at n = 150. The *F* test had slightly greater power than the *T* test for τ = 0.01 and 0.99 at $n < 150$ and equivalent power otherwise.

5. Example Application

Legendre et al. (1997) and Legendre and Legendre (1998:745-746) evaluated the contributions of spatial trend, physical habitat variables, and biotic interactions to bivalve mussel distribution and abundance in a New Zealand harbor. Physical habitat

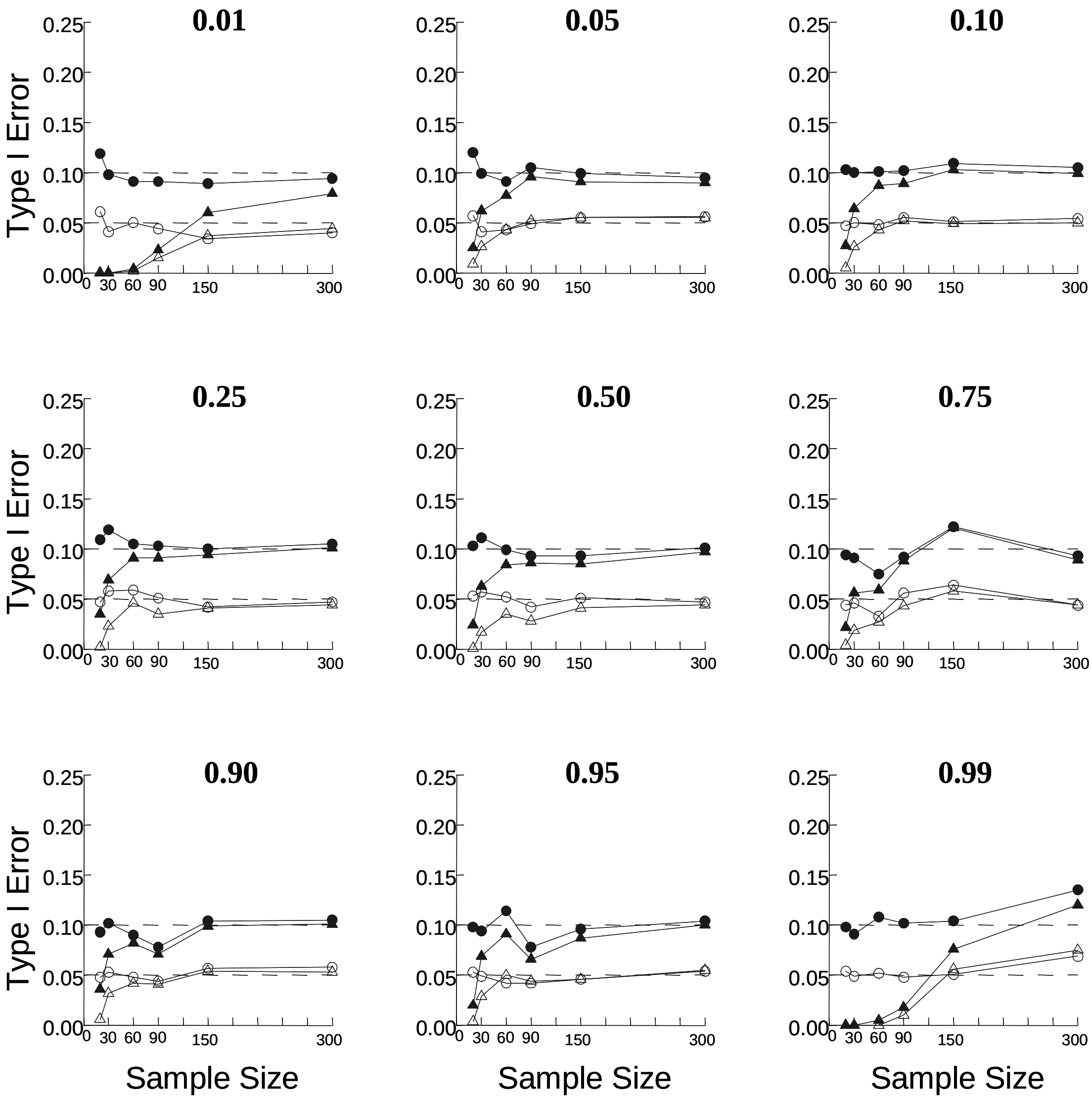


Figure 4.12. Estimated Type I error rates for α = 0.05 (open) and 0.10 (solid); for the *T* (triangles) and permutation *F* (circles) rankscore tests for H_0: $\beta_2(\tau) = \beta_3(\tau) =, \ldots, = \beta_{10}(\tau) = 0$ in the estimating model $y = \beta_0 X_0 + \beta_1 X_1 + \beta_2 LAT + \beta_3 LONG + \beta_4 LAT^2 + \beta_5 LONG^2 + \beta_6 LAT{\times}LONG + \beta_7 LAT^2{\times}LONG + \beta_8 LAT{\times}LONG^2 + \beta_9 LAT^3 + \beta_{10} LONG^3 + \varepsilon'$, for τ = 0.99, 0.95, 0.90, 0.75, 0.50, 0.25, 010, 0.05, and 0.01 for the N = 10,000 grid cells generated by the model in Figure 4.3; and for n = 20, 30, 60, 90, 150, and 300. 1,000 random samples were used for each n.

variables included sediment characteristics, bed elevation, and hydrodynamic measures likely to affect larval deposition, transport of juveniles, food supply, and feeding behavior. There were many strong correlations among the physical habitat variables considered. Biotic interactions considered adult-juvenile interactions by adding abundance of bivalves in different size classes to the models. Effects of a spatial trend surface, abundance of competitors, and habitat conditions were partitioned by considering nested sets of models in a linear least squares regression (Legendre et al. 1997), following procedures of Borcard et al. (1992). I explored relationships for one species, *Macomona liliana*, using similar procedures but estimated with quantile regression. The dependent variable was 22-23 January 1994 counts of *Macomona* >15 mm size class in 0.25-m^2 quadrats randomly located within 200 grid cells on a 250 m × 500 m area on the sandflat of Wiroa Island, Manukau Harbor, New Zealand.

Similar steps in modeling bivalve counts used by Legendre et al. (1997) were followed but several adjustments were made because I used regression quantile estimates and because I had a slightly different philosophy regarding model selection. Because the regression quantile estimates were intended to model heterogeneous variation in response distributions, I did not normalize bivalve counts by taking logarithms as did Legendre et al. (1997). When selecting polynomial terms to include in the final spatial trend surface model, I considered models with all linear terms; all linear and quadratic terms; and all linear, quadratic and cubic terms; which resulted in comparisons of three spatial trend models. I did not eliminate any individual monomial term from the set of linear, quadratic, or cubic polynomial terms.

I used $R^1(\tau)$ coefficients of determination (Koenker and Machado 1999) to compare fits of different regression quantile models across $\tau = 0.05 - 0.95$ by increments of 0.05. However, $R^1(\tau)$ like R^2 from least squares regression can only increase with increasing number of parameters and, thus, it was desirable to have a statistic that adjusts for inclusion of additional parameters relative to sample size. Therefore, I selected among models using a small sample size corrected version of the Akaike Information Criterion (AIC_c) developed by Hurvich and Tsai (1990) for the 0.50 regression quantile (i.e., least absolute deviation regression) and extended to other quantiles; $AIC_c(\tau) = 2n \times \ln(SAF(\tau)/n) + 2p(n/(n - p - 1))$, where $SAF(\tau)$ was the weighted sum of absolute deviations minimized in estimating the τth quantile regression with p parameters (including 1 for estimating σ). Appendix 4.1 describes computations for $R^1(\tau)$ and $AIC_c(\tau)$ and their justification. I computed differences ($\Delta AIC_c(\tau)$) between $AIC_c(\tau)$ for more complex models and the simplest model with just a constant (β_0) to facilitate comparisons among models in a fashion comparable to using coefficients of determination.

The modeling steps that Legendre et al. (1997) and I used were to (1) select an appropriate polynomial spatial trend surface model for bivalve counts; (2) select an appropriate model for bivalve counts as a function of the physical environmental variables; (3) test whether the spatial trend surface explained a significant fraction of additional variation given that the physical environmental variables were already in the model; (4) test whether the counts of competitors (bivalves in larger size classes) explained a significant fraction of additional variation given that the physical environmental variables were already in the model; and (5) test whether the spatial trend surface explained a significant fraction of additional variation given that the physical environmental and biological (abundance of competitors) variables were already in the model. Because *Macomona* >15 mm had no competitors by this protocol, steps (4) and (5) were not conducted for this size class. I present a condensed summary of results for counts of *Macomona* 0.5-2.5 mm where steps (4) and (5) were conducted.

Legendre et al. (1997) fit a spatial trend surface model first to determine whether there was any spatial structuring at the scale of the study plot associated with effects of ecological processes. However, I also consider the spatial trend surface as a potential surrogate for effects of unmeasured processes to be included in models after having accounted for effects associated with the measured variables.

5.1 Spatial Trend Surface

Plots of $R^1(\tau)$ coefficients of determination and $AIC_c(\tau)$ across $\tau = 0.05 - 0.95$ indicated that the linear + quadratic + cubic polynomial explained the greatest proportion of variation in counts of *Macomona* >15 mm and was the preferred trend surface model (Fig. 4.13). Trend surfaces plotted for the 0.90, 0.50, and 0.10 quantiles (Fig. 4.14) had similar variation along the northwest to southeast axis as the least squares regression surface estimated by Legendre et al. (1997), but the regression quantile estimates indicated greater variation associated with counts in the northwest corner and associated differences in rates of change in the spatial trend for different quantiles (note the trend surfaces in Figure 4.14 are not parallel in all regions). Regression quantile

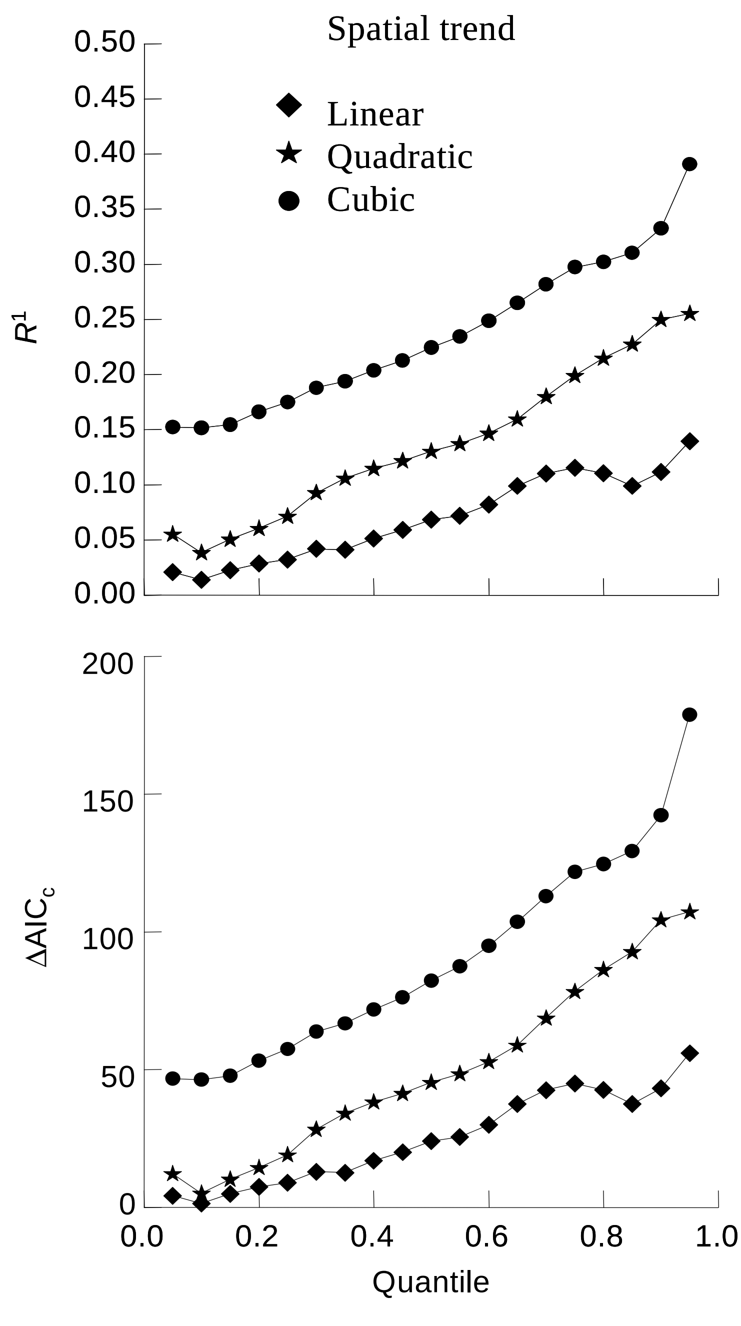


Figure 4.13. $R^1(\tau)$ coefficients of determination and differences in Akaike Information Criteria [$\Delta \mathrm{AIC_c}(\tau)$] for the linear (diamonds), linear + quadratic (stars), and linear + quadratic + cubic (circles) polynomial spatial trend surfaces for $\tau = 0.05$ to 0.95 (by increments of 0.05) regression quantiles of *Macomona liliana* >15 mm counts in 0.25-m^2 quadrats (n = 200), 22-23 January 1994, on the sandflat of Wiroa Island, Manukau Harbor, New Zealand (data from Legendre et al. 1997). All $\Delta \mathrm{AIC_c}(\tau)$ were computed by subtracting the $\mathrm{AIC_c}(\tau)$ for the model with just an intercept (β_0) from the $\mathrm{AIC_c}(\tau)$ for the polynomial models.

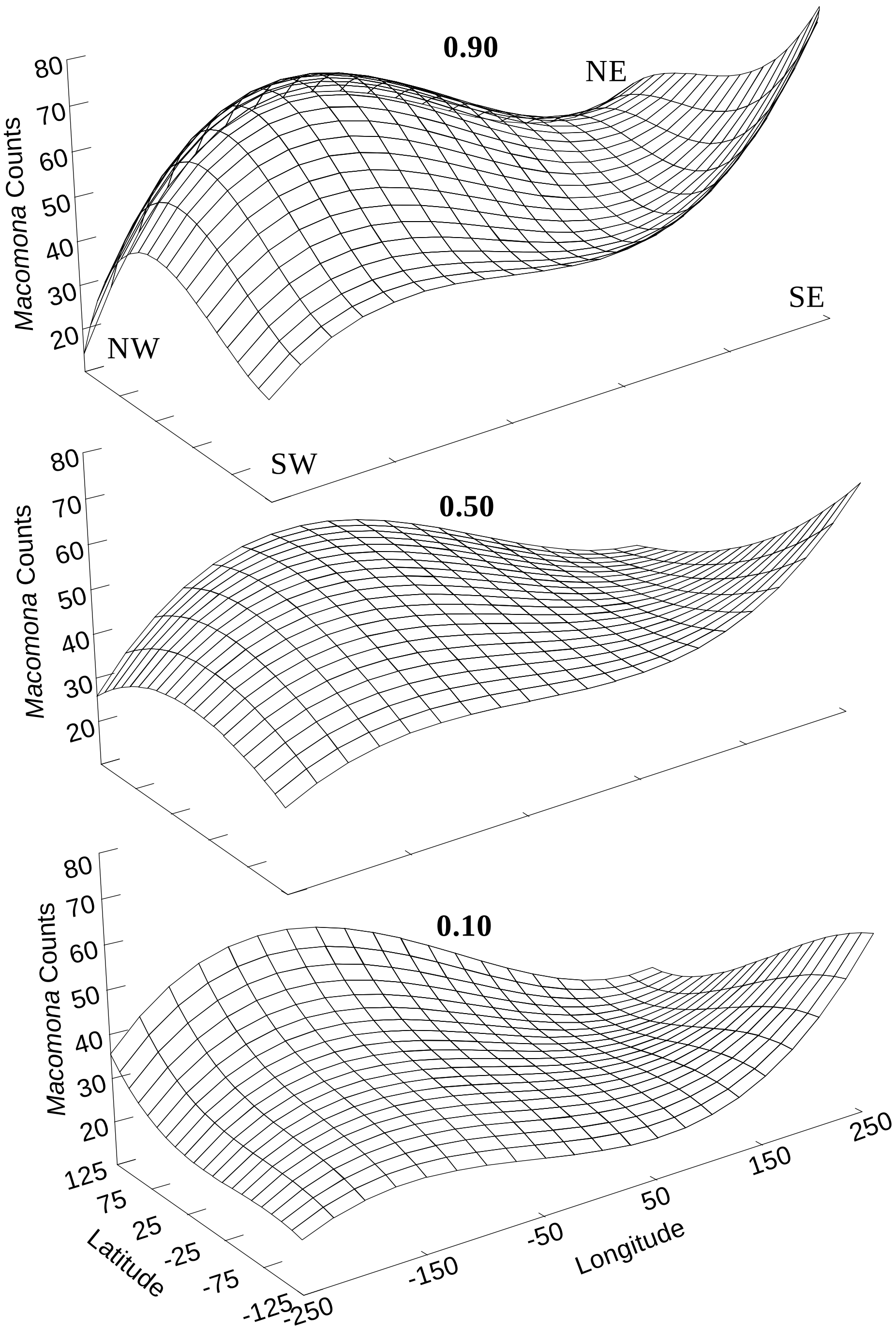


Figure 4.14. Cubic polynomial spatial trend surfaces for the 0.90, 0.50, and 0.10 regression quantiles of *Macomona liliana* >15 mm counts in 0.25-m^2 quadrats (n = 200), 22-23 January 1994, on the sandflat of Wiroa Island, Manukau Harbor, New Zealand. Latitude and longitude were centered to mean zero. View is from the southwest corner of the site. See text for estimated coefficients.

estimates established that variation in abundance and not just mean abundance of *Macomona* >15 mm had a spatial trend on the Wiroa sandflat. $R^1(\tau)$ coefficients of determination indicated substantially more variation explained at higher than at lower quantiles (Fig. 4.13). An ordinary least squares regression of *Macomona* >15 mm counts (not log transformed) on the same cubic polynomial had an $R^2 = 0.371$, which when returned to original units rather than squared units by $1 - (1 - R^2)^{0.5}$ (Ehrenberg 1975:233) indicated only 0.201 proportion of the variation in *Macomona* >15 mm counts was explained by the mean trend surface function. This was similar to variation explained by the central ($0.40 < \tau < 0.60$) regression quantile estimates but less than explained by higher quantiles (Fig. 4.13).

5.2 Physical Habitat

Legendre et al. (1997) found only two physical habitat variables explained any of the variation in ln mean counts of *Macomona* >15 mm: bed elevation (m) and percent of time the plot was covered by >20 cm of water during spring tide. These also were the only physical habitat variables that I found explained any of the variation in quantiles of large *Macomona*. However, these two variables were near perfectly linearly correlated ($r = -0.999$), which made physical sense because greater bed elevation is directly related to less water depth during high tides. I, therefore, chose to use only bed elevation in the physical habitat model. Legendre et al. (1997) used a cubic polynomial of bed elevation to model the nonlinear response of large *Macomona* counts (Fig 4.15). I initially considered this model too but also examined a simpler quadratic polynomial and compared models based on $R^1(\tau)$ and $\mathrm{AIC}_c(\tau)$. There was very little improvement

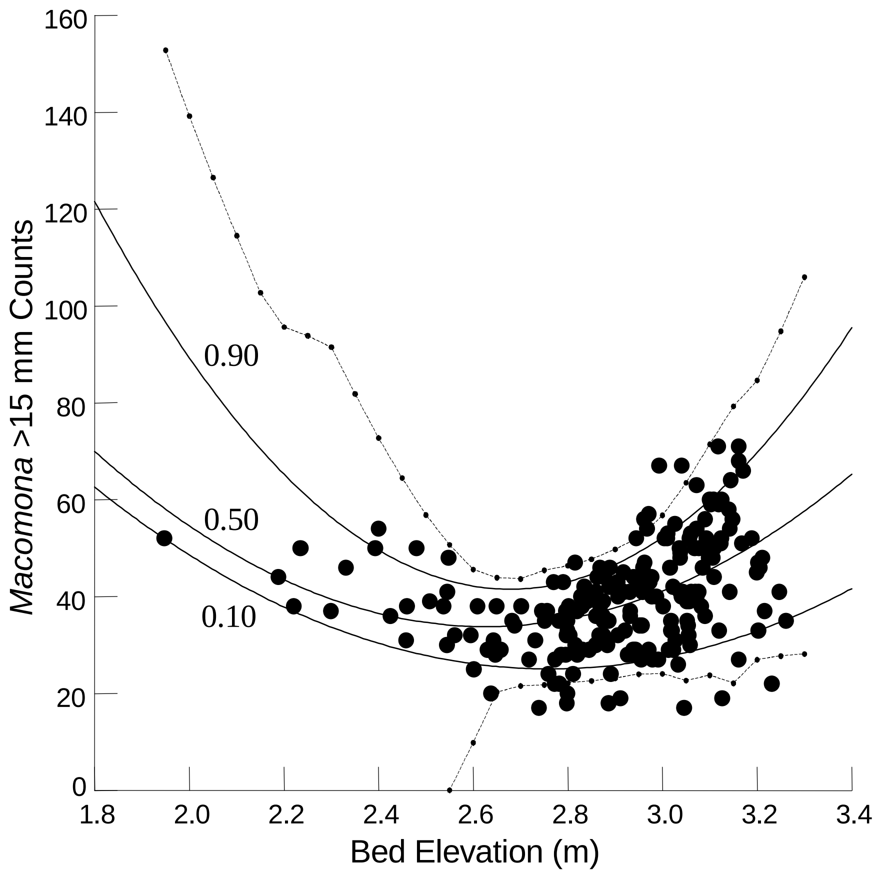


Figure 4.15. Counts of *Macomona liliana* >15 mm in 0.25-m^2 quadrats (n = 200), 22-23 January 1994, on the sandflat of Wiroa Island, Manukau Harbor, New Zealand, by bed elevation (m). Solid lines are 0.90, 0.50, and 0.10 regression quantile estimates of *Macomona* counts as a quadratic function of bed elevation. Lines with small dots connect upper and lower Working-Hotelling 80% simultaneous confidence intervals for predicted 0.90 (upper) and 0.10 (lower) regression quantiles at 28 selected values of bed elevation.

in coefficients of determination by going to the cubic compared to the quadratic polynomial (Fig. 4.16). Differences in $\Delta AIC_c(\tau)$ supported use of the cubic polynomial of bed elevation only for 0.80 - 0.85 quantiles. An examination of the cubic polynomial model of bed elevation suggested that regression quantile fits that were better with the cubic term were greatly influenced by the outlying minimum elevation value of 1.95 m. Removing this influential value and estimating quadratic and cubic polynomial models and associated fit and model selection statistics again indicated even less support for including the cubic bed elevation term.

The quadratic response of large *Macomona* to bed elevation captured the higher counts at lower and higher bed elevations and increasing variation in counts at higher elevations (Fig. 4.15). Unweighted estimates and 90% confidence intervals for linear (b_1) and quadratic (b_2) terms indicated increasingly negative linear and increasingly positive quadratic terms with increasing τ above 0.50 (Fig. 4.17). Nonlinear changes in large *Macomona* with respect to bed elevation were greater for higher quantiles. Although heterogeneity in counts across bed elevation was not extreme, I constructed weighted regression quantile estimates for τ = 0.05 - 0.95 by increments of 0.05, where weights were estimated separately for each individual quantile with a variant of the bandwidth approach used by Koenker and Machado (1999). I used the Hall and Sheather (1988) bandwidth selection rule recommended by Koenker and Machado (1999) but did not use their approach of taking differences between estimates for the highest and lowest quantile within the bandwidth. Instead, weights were computed by taking the average pairwise difference between all unweighted quantile estimates for

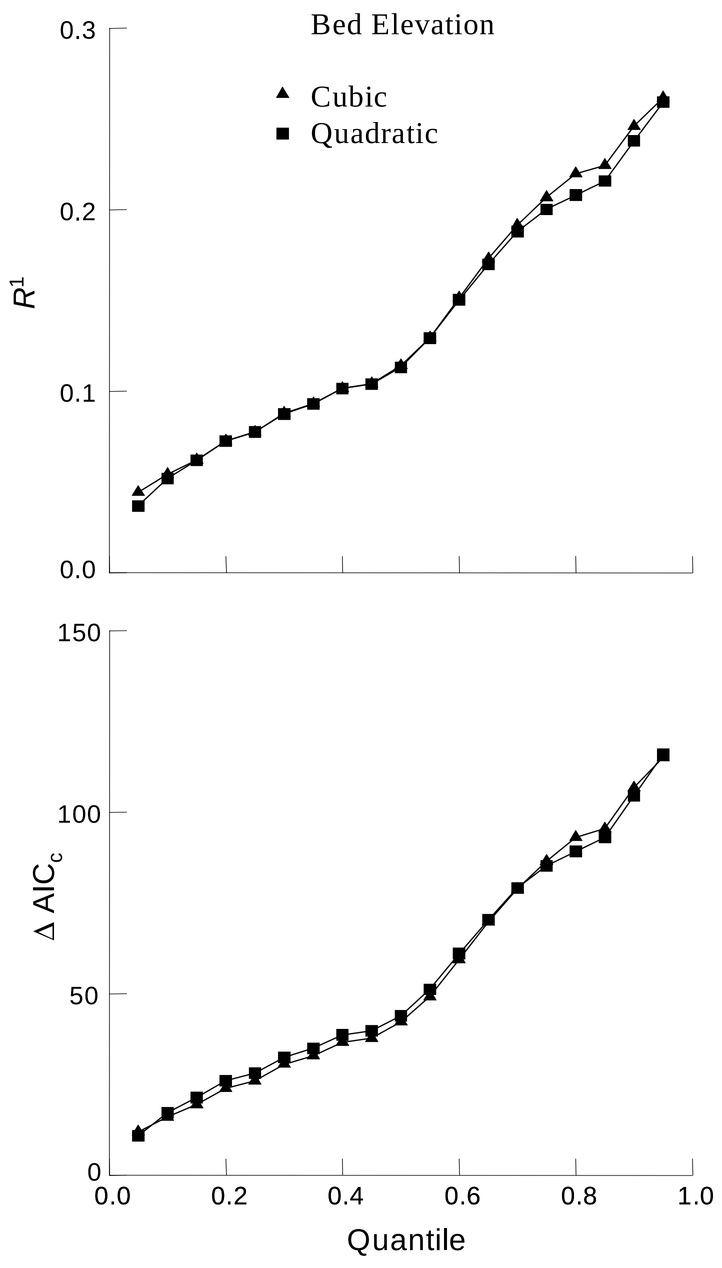


Figure 4.16. $R^1(\tau)$ coefficients of determination and differences in Akaike Information Criteria [$\Delta AIC_c(\tau)$] for the linear + quadratic (squares) and linear + quadratic + cubic (triangles) functions of bed elevation (m) for τ = 0.05 to 0.95 (by increments of 0.05) regression quantiles of *Macomona liliana* >15 mm counts in 0.25-m^2 quadrats (n = 200), 22-23 January 1994, on the sandflat of Wiroa Island, Manukau Harbor, New Zealand. All $\Delta AIC_c(\tau)$ were computed by subtracting the $AIC_c(\tau)$ for the model with just an intercept (β_0) from the $AIC_c(\tau)$ for the polynomial models.

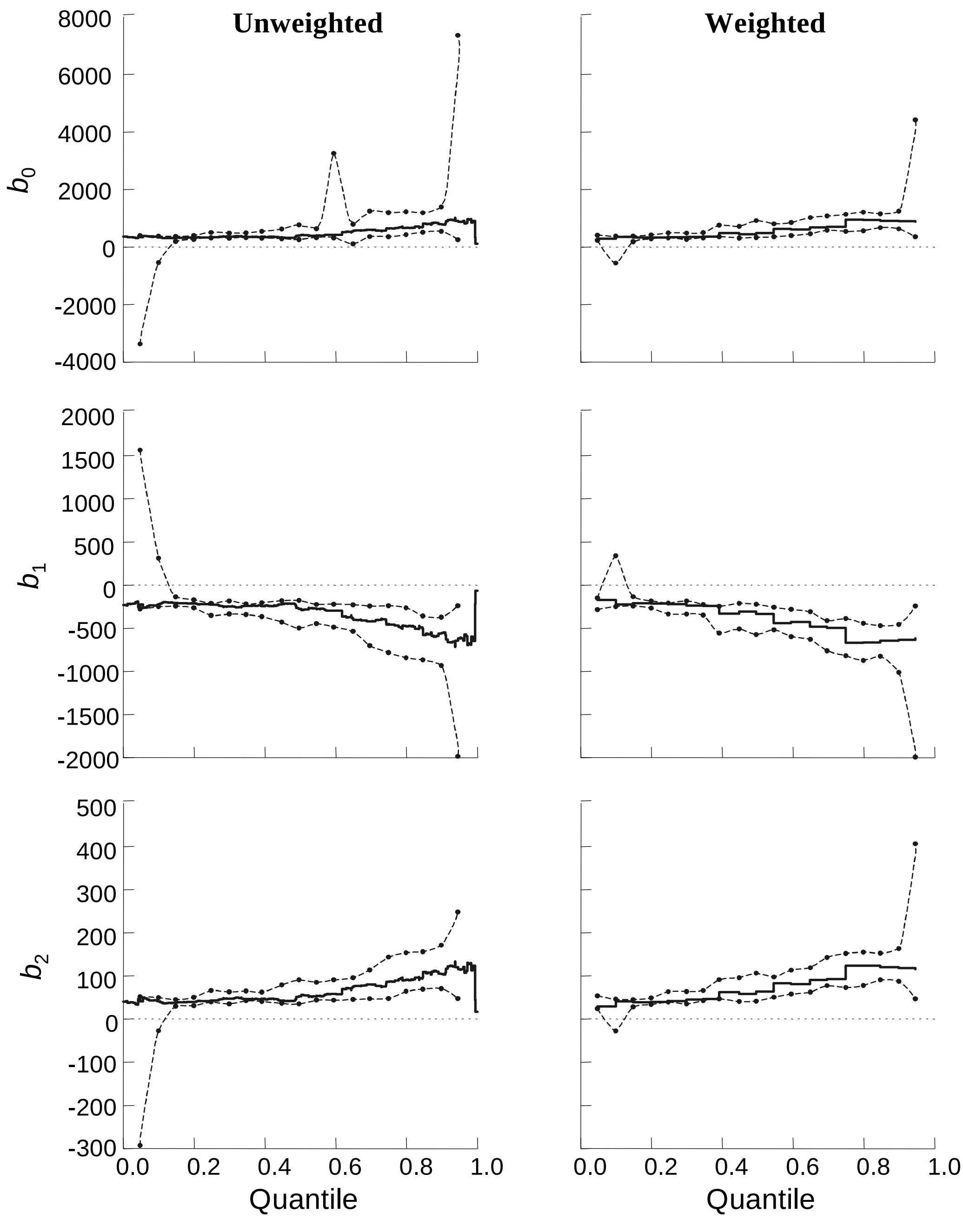


Figure 4.17. Estimates for intercept [$b_0(\tau)$], linear [$b_1(\tau)$], and quadratic [$b_2(\tau)$] terms for regression quantiles of *Macomona liliana* >15 mm counts in 0.25-m^2 quadrats (n = 200), 22-23 January 1994, on the sandflat of Wiroa Island, Manukau Harbor, New Zealand, as a quadratic function of bed elevation (m) for both unweighted and weighted models. Solid lines are step functions of parameter estimates by quantiles (τ), all for unweighted estimates and for τ = 0.05 - 0.95 by increments of 0.05 for weighted estimates. Dashed lines connect pointwise 90% confidence intervals based on inverting the T rankscore tests for τ = 0.05 - 0.95 by increments of 0.05.

$b_0(\tau)$, $b_1(\tau)$, and $b_2(\tau)$ within the interval $\tau \pm h(\tau)$, where $h(\tau)$ was the bandwidth for a specified quantile. This reduced the number of negative weights due to crossing of regression quantile estimates at extreme regions of the design matrix that occurred with the method used by Koenker and Machado (1999). Still, small constants had to be added to the average pairwise differences for $b_0(\tau)$ to assure positive weights for a couple of quantiles. Example calculations for the weights are in Appendix 4.2.

Weighted estimates for the quadratic polynomial terms of bed elevation followed a similar pattern of changes with quantiles as the unweighted estimates, although weighted estimates smoothed over a little detail because they were only done for 19 increments of τ between 0.05 and 0.95 (Fig. 4.17). The 90% confidence intervals for the weighted estimates were slightly narrower than those for the unweighted estimates at most higher quantiles. One unusually large upper endpoint of the interval for the unweighted estimate $b_0(0.60)$ was eliminated in the weighted estimate. The overall pattern and inference for weighted estimates did not differ substantially from those for unweighted estimates. This was consistent with the moderate amount of heterogeneity in *Macomona* >15 mm counts across bed elevation (Fig. 4.15).

A simultaneous 80% prediction interval on the central 80% of *Macomona* >15 mm as a function of bed elevation was estimated by constructing simultaneous confidence intervals for the 0.10 and 0.90 regression quantile estimates at 25 values of bed elevation between 2.10 and 3.30 m (Fig. 4.15). The simultaneous prediction intervals emulated the Working-Hotelling simultaneous confidence intervals (Neter et al. 1996:234) for estimates $b_0(\tau)$ with bed elevation shifted by the 25 selected values for prediction; the zero intercept was shifted to correspond to the selected values of bed elevation (Chapter 2). Two-sided intervals were constructed by inverting the weighted quantile rankscore test with an $\alpha = 0.0316$ [1 - prob. $F((3 \times F(0.80, 3, 197)), 1, 197)$] with the upper part of the interval for $b_0(0.90)$ and the lower part of the interval for $b_0(0.10)$. The interval displayed in Figure 4.15 was, thus, a statement about where the central 80% of *Macomona* >15 mm would be expected to occur with respect to bed elevation in 80% of repeated random samples. Note that lower intervals for $\tau = 0.10$ and bed elevation $\leq$2.5 m extended below zero counts, a nonsensical value. The wide intervals here were likely due to fewer observations with bed elevations $\leq$2.5 m and because the lower portion of the distribution of large *Macomona* counts were not as well defined as the upper part. Slight irregularities in the upper and lower simultaneous confidence intervals should not be over interpreted as they likely occurred due to the vagaries of interpolating between discrete probabilities associated with the rankscore test statistics (Chapter 2). Use of a more stringent confidence level such as 90% required smaller individual α's that resulted in intervals with greater irregularities.

5.3 Physical Habitat Plus Spatial Trend

Including the cubic polynomial spatial trend surface indicated that there was additional variation in *Macomona* >15 mm abundance that was spatially structured after accounting for effects of bed elevation (Fig. 4.18). Changes in $\Delta AIC_c(\tau)$ clearly supported the model with bed elevation and the spatial trend surface over the model with just bed elevation (Fig.4.18), but the relative sampling frequency probabilities for

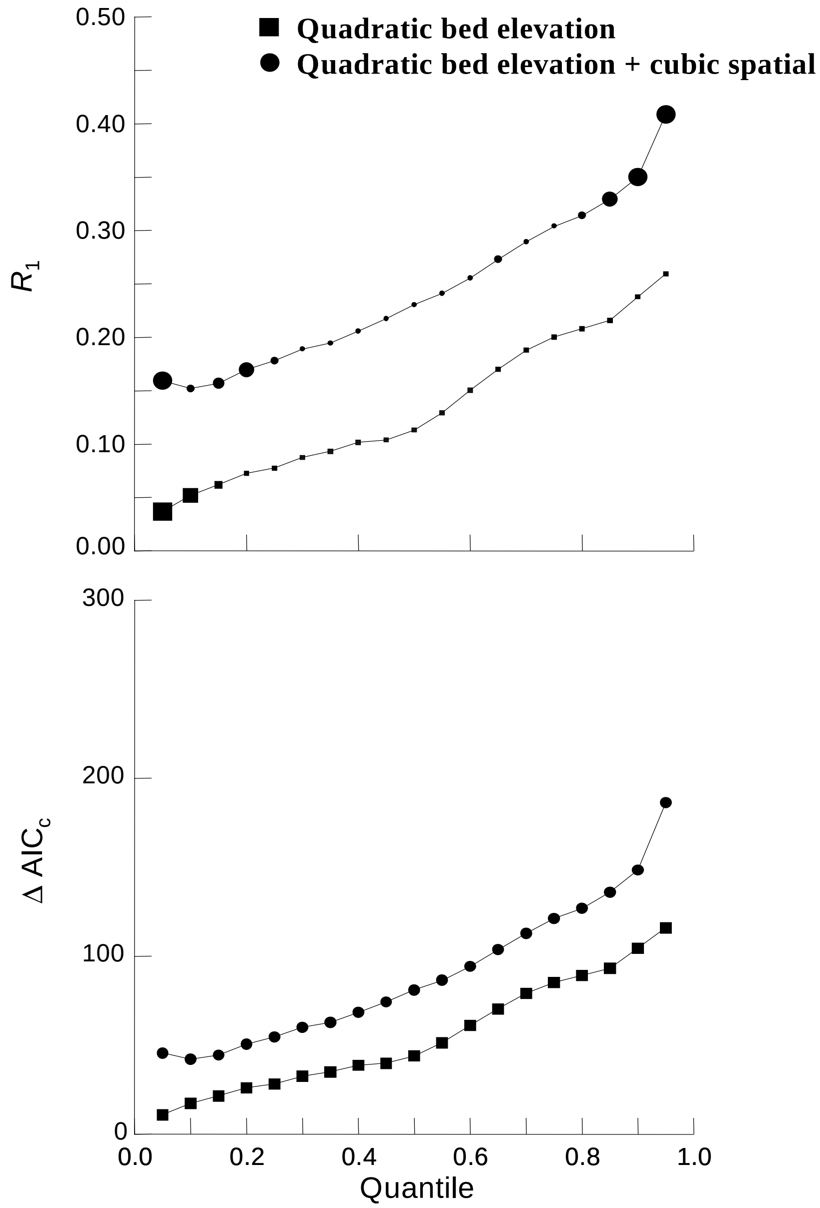


Figure 4.18. $R^1(\tau)$ coefficients of determination and differences in Akaike Information Criteria [$\Delta\text{AIC}_c(\tau)$] for the physical habitat model as a quadratic function of bed elevation (squares) and for the physical habitat + cubic polynomial spatial trend surface (circles) for $\tau = 0.05$ to 0.95 (by increments of 0.05) regression quantiles of *Macomona liliana* >15 mm counts in 0.25-m^2 quadrats ($n = 200$), 22-23 January 1994, on the sandflat of Wiroa Island, Manukau Harbor, New Zealand. All $\Delta\text{AIC}_c(\tau)$ were computed as in Figure 4.13. Symbol sizes for $R^1(\tau)$ are proportional in size to $P(\text{H}_0$: linear + quadratic bed elevation parameters = 0) for physical habitat model (squares) and $P(\text{H}_0$: cubic polynomial spatial trend surface parameters = 0|linear + quadratic bed elevation parameters) (circles) from the permutation F rankscore test; largest symbol is for $P \geq 0.20$ and smallest is for $P < 0.01$. P for T rankscore tests were similar.

lower ($\tau \leq 0.20$) and higher ($\tau \geq 0.85$) quantiles indicated the joint effects of the polynomial spatial coefficients were not different from zero. Because bed elevation itself was spatially structured (Legendre et al. 1997), estimated effects of bed elevation after adjusting for spatial trend were attenuated and reversed in sign because they were confounded with other processes associated with spatial trend (Fig. 4.19). In this model, 90% confidence intervals for bed elevation included zero for all quantiles (Fig. 4.19). Here only unweighted estimates were used, as the previous analysis on bed elevation suggested effects of heterogeneity were not great enough for weighted confidence intervals to differ substantially from unweighted ones. The cubic polynomial spatial trend surface model explained nearly as much variation as the model that included quadratic bed elevation and cubic spatial trend (compare Figs. 4.13 and 4.18). The cubic polynomial spatial trend surface given effects of bed elevation retained most of the northwest to southeast variation estimated by the spatial trend surface alone, except that some of the variation in the southeast corner was attenuated (compare Figs. 4.19 and 4.14). When the $R^2 = 0.33$ for the physical habitat and spatial trend of ln mean abundance of large *Macomona* estimated by Legendre et al. (1997) is converted into original units [$0.18 = 1 - (1 - 0.33)^{0.5}$, Ehrenberg 1975:233), it is obvious that the regression quantile estimates (not in ln scale) for a similar model explained a greater proportion of variation across $\tau \geq 0.40$ (Fig. 4.18).

5.4 Summary for Macomona 0.5-2.5 mm

Abundance of *Macomona* 0.5-2.5 mm was determined for 3 cores totaling 0.04-m^2 within the 0.25-m^2 primary quadrats (Legendre et al. 1997). Counts ranged from 0 to

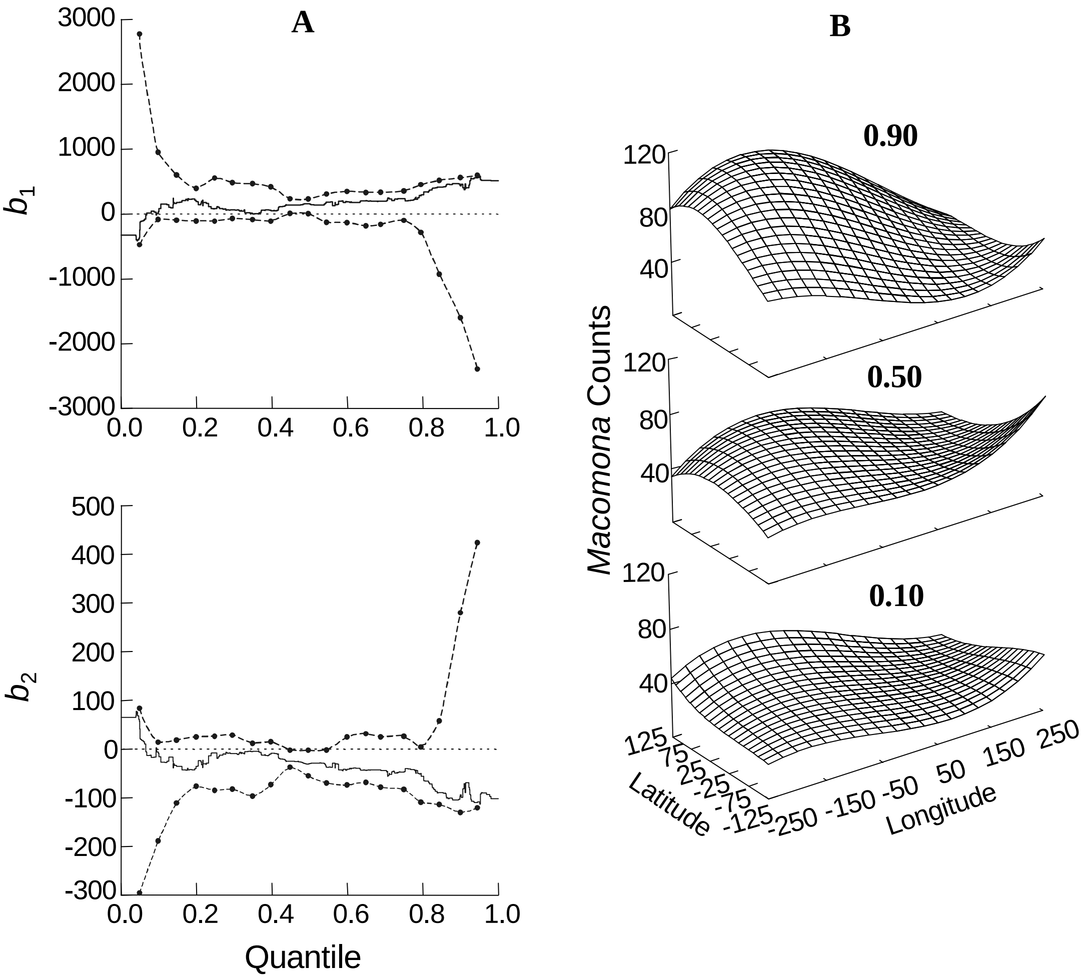


Figure 4.19. (A) Estimates for linear [$b_1(\tau)$] and quadratic [$b_2(\tau)$] terms for regression quantiles of *Macomona liliana* >15 mm counts in 0.25-m^2 quadrats (n = 200), 22-23 January 1994, on the sandflat of Wiroa Island, Manukau Harbor, New Zealand, as a quadratic function of bed elevation (m) after adjusting for the cubic polynomial spatial trend surface. Solid lines are step functions of parameter estimates by quantiles (τ) and dashed lines connect pointwise 90% confidence intervals based on inverting the T rankscore tests for τ = 0.05 - 0.95 by increments of 0.05. (B) The 0.90, 0.50, and 0.10 cubic polynomial spatial trend surfaces after adjusting for the quadratic function of bed elevation at the mean value of 2.9 m. View is from the southwest.

12. Following the previous steps outlined including steps 4 and 5 that added effects of biological competitors, I estimated regression quantiles for counts (not ln counts) of small *Macomona* as a function of 4 physical habitat variables, 4 physical habitat variables plus quadratic spatial trend, 4 physical habitat variables plus abundance of 3 biological competitors, and 4 physical habitat variables plus abundance of 3 biological competitors plus quadratic spatial trend (Fig. 4.20). Comparisons of $R^1(\tau)$ and $AIC_c(\tau)$ selected the same variables used by Legendre et al. (1997) for their mean ln abundance model, with several minor changes. I retained all terms of the quadratic polynomial spatial trend surface. I did not consider including both bed elevation and proportion of the time the plot was covered by >20 cm water during spring tide because they were near perfectly correlated. The 4 physical habitat variables used were a linear function of bed elevation (m), shell hash (g, not square root transformation), peak ebb-tide bed shear stress ($N \cdot m^{-2}$) and peak flood-tide bed shear stress ($N \cdot m^{-2}$, not log transformed). Transformations used by Legendre et al. (1997) for shell hash and flood-tide shear stress were not required for obtaining reasonable regression quantile estimates. The 3 variables measuring potential biological competitors included counts (not ln transformed) of *Macomona* 2.5-4.0 mm, *Austrovenus stutchburyi* 0.5-2.5 mm, and *Austrovenus* 2.5-4.0 mm.

Low proportion of variation explained by all models at lower quantiles (τ <0.20) was due to the proportion of zero counts that were modeled as well with just a constant (Fig. 4.20). After $R^1(\tau)$ increased up to $\tau = 0.35$, there was a fairly constant amount of variation explained by all higher quantiles regardless of the model. This was

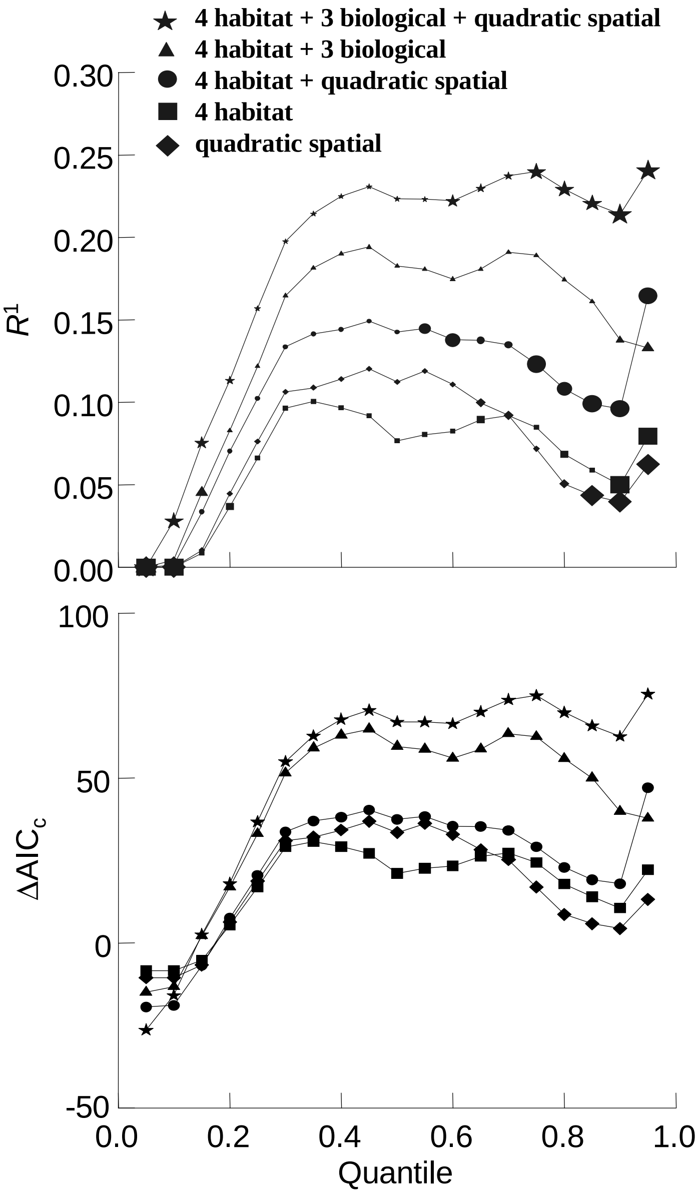


Figure 4.20. $R^1(\tau)$ and $\Delta AIC_c(\tau)$ for the physical habitat model as a linear function of bed elevation (m), shell hash (g), ebb-tide shear stress ($N \cdot m^{-2}$), and flood-tide shear stress ($N \cdot m^{-2}$) (squares); for the quadratic polynomial spatial trend surface (diamonds); for the physical habitat + quadratic spatial trend surface (circles); for the physical habitat + biological competitors as a linear function of *Austrovenus* 0.5-2.5 mm, *Austrovenus* 2.5-4.0 mm, and *Macomona* 2.5-4.0 mm counts (triangles); and for the physical habitat + biological competitors + quadratic spatial trend surface (stars) for $\tau = 0.05$ to 0.95 (by increments of 0.05) regression quantiles of *Macomona liliana* 0.5-2.5 mm counts in 0.04-m^2 quadrats ($n = 185$), 22-23 January 1994, on the sandflat of Wiroa Island, Manukau Harbor, New Zealand. All $\Delta AIC_c(\tau)$ were computed as in Figure 4.13. Symbol sizes for $R^1(\tau)$ are proportional in size to $P(H_0$: 4 physical habitat parameters = 0) (squares); $P(H_0$: quadratic spatial trend surface parameters = 0) (diamonds); $P(H_0$: quadratic spatial trend surface parameters = 0|4 physical habitat parameters) (circles); $P(H_0$: 3 biological competitor parameters = 0|4 physical habitat parameters) (triangles); $P(H_0$: quadratic spatial trend surface parameters = 0|4 physical habitat + 3 biological competitor parameters) (stars) from the permutation *F* rankscore test; largest symbol is for $P \geq 0.20$ and smallest is for $P < 0.01$. *P* for *T* rankscore tests were similar.

consistent with the low range of variation in counts of *Macomona* 0.5-2.5 mm. Estimated effects for physical habitat variables were increasingly negative for shell hash with increasing τ, positive for flood-tide shear stress and increasing for τ >0.80, increasingly positive for ebb-tide shear stress up to $\tau = 0.80$ then increasingly negative for τ >0.80, and positive for bed elevation and increasing for τ >0.60. Estimated effects for abundance of biological competitors were positive and increasing with τ for *Macomona* 2.5-4.0 mm, *Austrovenus* 2.5-4.0 mm, and *Austrovenus* 0.5-2.5 mm. My estimated quantile regression effects of biological competitors were positive as were effects estimated by Legendre et al. (1997), suggesting that they were not measuring competition for resources but may have been surrogates for unmeasured physical factors that affected deposition and settlement of juveniles similarly across taxa.

Differences in quantile regression models of abundance parallel those of the mean ln abundance regression models estimated by Legendre et al. (1997) except that the regression quantile estimates suggested that variation due to quadratic spatial trend existed even after including the 4 physical habitat and 3 biological competitor variables. However, relative sampling frequency probabilities associated with $\tau \geq 0.75$ suggested the joint effects of quadratic polynomial terms were not different from zero when considering spatial trend added to either the model with just physical habitat or the model with physical habitat and biological competitor abundance (Fig. 4.20). The quadratic spatial trend surface after adjusting for physical habitat and abundance of potential biological competitors had its main axis of change from the southeast to northwest with greater variation in abundance in the southeastern part of the sandflat.

Transforming $R^2 = 0.41$ estimated by Legendre et al. (1997) for the physical habitat, biological competitors, and quadratic spatial trend model to original units [$0.23 = 1 - (1 - 0.41)^{0.5}$] indicated that the amount of variation explained by the mean ln abundance model was comparable to that explained for most quantiles ($\tau > 0.35$) of abundance with a similar model (Fig. 4.20). However, Legendre et al. (1997) concluded that quadratic spatial trend did not explain a significant proportion of variance in mean ln abundance after accounting for physical habitat and biological competitors, whereas my estimates indicated that it did for most quantiles ($\tau < 0.75$) of abundance.

6. Discussion

The example simulations demonstrated how heterogeneous and nonlinear relations in habitat models can easily arise from confounding with some important but unmeasured processes. More complicated arguments are not required to explain why heterogeneity and nonlinearities are so common in statistical models of animal responses to their habitat resources. Although the dimensions of the measured habitat variables (X_1) and the unmeasured limiting factors (X_2) were kept to single variables for my simulation purposes, it is reasonable to extend interpretation of these simulation results to greater dimensions by thinking of X_1 and X_2 as being the composite additive effect of >2 variables. More complicated interactions involving both interference (- coefficients) and facilitation (+ coefficients) interactions should be additive and may have attenuated overall effect depending on the magnitude of the separate effects. My presentation of the simulations focused on confounding with unmeasured variables not related to habitat resources. It is easy to extend my results and interpretations to situations where confounding occurs with some important habitat resources that were not measured and included in the model used for estimation.

The philosophy embodied in my simulations reflect a view that most ecological relations have an appearance of randomness not because they are inherently random but because we are always estimating them with incomplete information (Regan et al. 2002). As long as random variation induced by missing information is small and homogeneous, conventional regression estimation procedures (e.g., least squares) may provide useful, reasonable estimates of conditional relationships. When missing information is for processes of substantial importance to an organism, it is reasonable to expect large, heterogeneous random variation and estimates with hidden bias. While all organisms are dependent on some suite of resources obtained from their habitat, at many times and locations other factors may actually exert more influence on organism growth, survival, reproduction, and dispersal, causing a perceived disconnection between the organism and the requisite habitat resources. Garshelis (2000) and Morrison (2001) both have argued for improving our knowledge of animal habitat relations by focusing modeling efforts on more specifically defined resources and relating them to demographic parameters such as survival and reproductive rates that ultimately contribute to differences in abundance. These are reasonable suggestions. But, neither a more focused definition of what constitutes a habitat resource or measuring alternative demographic parameters will eliminate issues of hidden bias due to confounding between measured habitat covariates and unmeasured ones associated with other processes.

Inference procedures based on rankscores for regression quantile estimates provided valid intervals reflecting the sampling distribution of parameter estimates for the measured habitat processes, but the parameters clearly were biased relative to the parameters generating the responses. In applications, the degree of hidden bias will be greater or lesser for different quantiles depending on the nonestimable interaction effects and unknown error distributions. If it is possible to rule out certain types of interaction effects (e.g., facilitation) with unmeasured processes, then we might profitably focus estimation and inference procedures for quantile regression at one end of the probability distribution (e.g., upper quantiles). In the absence of such knowledge, it would appear prudent to obtain estimates and confidence intervals across the entire interval of quantiles that provide reliable estimates (e.g., 0.05 - 0.95). I encourage the use of prediction intervals, and especially simultaneous prediction intervals or tolerance intervals, as a strong antidote to overzealous expectations that any habitat model can provide precise predictions. It simply is not reasonable to expect that habitat models should provide very precise predictions when they exclude many other important processes. However, this does not imply that useful predictions are impossible with habitat models, especially for management or conservation purposes. We simply have to use better procedures for characterizing intervals of response such as those presented here and have more realistic expectations about predicting changes in populations due to changes in habitat, a process that interacts with other processes that we often barely understand or know how to measure.

Prediction and tolerance intervals provide confidence statements related to individual observational units (Vardeman 1992). These were areal plots in my simulations and example application. Prediction and tolerance intervals based on my regression quantiles are semiparametric in nature because they don't assume a specific parametric form for the unknown error distribution. They are likely to be much more informative for characterizing the real uncertainty in habitat models and for making predictions useful for management or conservation purposes than confidence intervals on mean rate parameters or responses. We never observe a mean response across observational units, only responses for individual units. Management or conservation actions are implemented on individual units of area in a landscape. Estimating an interval of responses likely for these individual areas provides more knowledge about what outcomes might be realized than is provided by confidence intervals associated with mean responses or any other individual parameter. Scientific attempts to improve predictions from habitat models are based on measuring observed outcomes relative to predictions necessarily obtained from measurements on individual units of area. Understanding the contexts where habitat models fail or succeed as predictors of population change can only be gained by considering contextual information for individual units of area on landscapes.

My simulation results demonstrated that heterogeneity that arises due to confounding between measured and unmeasured variables often will not be a simple location-scale form. In this situation, weighted regression quantile estimates and rankscore tests require estimating weights that are based on changes in a local interval of quantiles around a specific quantile rather than globally applied across all quantiles. I used a minor modification of bandwidth estimation procedures developed by Hall and Sheather (1988) as extended to regression quantiles by Koenker and Machado (1999). Although adequate, there clearly is room for improvement in these procedures, including automating their computation in the necessary software.

My use of ΔAIC_c for model selection with the bivalve data was an attempt to extend Hurvich and Tsai (1990) procedures for median regression (0.5 quantile) to other quantiles. The fact that some large ΔAIC_c between models at high and low quantiles were associated with sampling distributions of parameter estimates that did not differ from zero was a little disconcerting. This may reflect a failure of this extension of AIC_c, that the sampling distribution of estimates is not well represented in information criteria like AIC_c, or that I extended estimates and inferences too far into the extreme quantiles for them to be reliable. Machado (1993) discussed extension of Schwarz information criterion (SIC) to robust *M*-estimates, including median regression, for linear models. Additional research on application of information criteria to regression quantile model selection is clearly warranted.

Use of cubic polynomials of location coordinates to estimate spatial trend surfaces provided a reasonable method for modeling larger scale spatial gradients of responses (Legendre et al. 1997) that are of most interest for models generalizing animal responses to habitat. Spatial trend surfaces provided an indication of spatial variation in organism response which would suggest effects of some relevant ecological processes (Legendre et al. 1997) and provided a method for accounting for some of the variation due to unmeasured processes that were spatially structured. However, it is important to remember that gradients in space offer no ecological interpretation per se (Legendre et al. 1997). It is in fact possible to defeat the entire purpose of developing general habitat relationships by over reliance on modeling spatial structure. Consider the models of *Macomona* >15 mm as a function of bed elevation and spatial structure. There was more variation in large *Macomona* abundance explained by the spatial trend surface alone than by the nonlinear bed elevation model. A parsimonious model that explained most variation with fewest parameters would be the cubic spatial trend surface model. Yet, this model of bivalve counts based on spatial gradients on one sandflat has little chance of generalizing to other locations because it includes no information on ecological processes. The cubic spatial trend does suggest that spatially structured processes are operating within the scale of the sampled 250 m × 500 m area (Legendre et al. 1997). There is greater potential for generalizing the bed elevation relationship to other locations to the extent that bed elevation is related to hydrodynamic processes affecting settlement, feeding, and survival of bivalves.

An extreme form of spatial modeling that defeats the generality desired in most habitat models is including indicator variables for different geographic locations. This model structure allows for different relationships with the habitat covariates for every geographic location, i.e., unequal slopes and intercepts in a separate regression model. This might be justified based on statistical measures for regression model selection, fit, or hypothesis testing. Yet, separate relationships for each geographic location would seem to completely defeat our desire for developing general relationships in ecology.

With quantile regression, it is possible to have different geographic locations associated with different quantiles of one common model, e.g., Dunham et al. 2002. The difference from a statistical standpoint is whether to assume one common probability model with different locations (contexts) associated with different portions of the probability distribution (quantiles), or whether to assume separate probability models associated with each location. A desire to generalize to other places and times with habitat relationship models suggests that the common probability model might be more informative.

Acknowledgments

Jon D. Richards provided programming support for simulations. J. E. Roelle and J. W. Terrell reviewed drafts of the manuscript. K. P. Burnham reviewed the justification and computations for AIC_c. P. Legendre generously provided the bivalve mussel data and suggestions regarding the double permutation scheme.

Literature Cited

Bi, H., G. Wan, and N. D. Turvey. 2000. Estimating the self-thinning boundary line as a density-dependent stochastic biomass frontier. Ecology **81**:1477-1483.

Borcard, D., P. Legendre, and P. Drapeau. 1992. Partialling out the spatial component of ecological variation. Ecology **73**:1045-1055.

Cade, B. S., and Q. Guo. 2000. Estimating effects of constraints on plant performance with regression quantiles. Oikos **91**: 245-254.

Cade, B. S., and J. D. Richards. 1996. Permutation tests for least absolute deviation regression. Biometrics **52**:886-902.

Cade, B. S., J. W. Terrell, and R. L. Schroeder. 1999. Estimating effects of limiting factors with regression quantiles. Ecology **80**:311-323.

Dunham, J. B., B. S. Cade, and J. W. Terrell. 2002. Influences of spatial and temporal variation on fish-habitat relationships defined by regression quantiles. Transactions of the American Fisheries Society **131**: 86-98.

Eastwood, P. D., G. J. Meaden, and A. Grioche. 2001. Modeling spatial variations in spawning habitat suitability for the sole *Solea solea* using regression quantiles and GIS procedures. Marine Ecology Progress Series **224**: 251-266.

Ehrenberg, A. S. C. 1975. Data reduction: analyzing and interpreting statistical data. Wiley, New York.

Fausch, K. D., C. L. Hawks, and M. G. Parsons. 1988. Models that predict standing crop of stream fish from habitat variables (1950-85). U. S. Forest Service General Technical Report PNW-GTR-213.

Garshelis, D. L. 2000. Delusions in habitat evaluation: measuring use, selection, and importance. Pages 111-164 *in* L. Boitani and T. K. Fuller, eds. Research Techniques in Animal Ecology. Columbia University Press, New York.

Haire, S. L., C. E. Bock, B. S. Cade, and B. C. Bennett. 2000. The role of landscape and habitat characteristics in limiting abundance of grassland nesting songbirds in an urban open space. Landscape and Urban Planning **48**(1-2):65-82.

Hall, P., and S. Sheather. 1988. On the distribution of a studentized quantile. Journal of the Royal Statistical Society, Ser. B **50**:381-391.

Hurvich, C. M., and C-L. Tsai. 1990. Model selection for least absolute deviations regression in small samples. Statistics and Probability Letters **9**:259-265.

Huston, M. A. 2002. Introductory essay: Critical issues for improving predictions. Pages 7-21 *in* J. M. Scott et al., editors. Predicting species occurrences: issues of accuracy and scale. Island Press, Covelo, California, USA.

Kaiser, M. S., P. L. Speckman, and J. R. Jones. 1994. Statistical models for limiting nutrient relations in inland waters. Journal of the American Statistical Association **89**:410-423.

Koenker, R., and J. A. F. Machado. 1999. Goodness of fit and related inference processes for quantile regression. Journal of the American Statistical Association **94**:1296-1310.

Legendre, P., and L. Legendre. 1998. Numerical ecology. Second English edition. Elsevier Science, Amsterdam, The Netherlands.

Legendre, P., S. F. Thrush, V. J. Cummings, P. K. Dayton, J. Grant, J. E. Hewitt, A. H. Hines, B. H. McArdle, R. D. Pridmore, D. C. Schneider, S. J. Turner, R. B., Whitlatch, and M. R. Wilkinson. 1997. Spatial structure of bivalves in a sandflat: Scale and generating processes. Journal of Experimental Marine Biology and Ecology **216**:99-128.

Lichstein, J. W., T. R. Simons, S. A. Shriner, and K. E. Franzreb. 2002. Spatial autocorrelation and autoregressive models in ecology. Ecological Monographs **72**:445-463.

Machado, J. A. F. 1993. Robust model selection and *M*-estimation. Econometric Theory **9**: 478-493.

McQuarrie, A. D. R., and C-L. Tsai. 1998. Regression and time series model selection. World Scientific Publishing Co. Pte. Ltd. Singapore. 455pp.

Morrison, M. L. 2001. A proposed research emphasis to overcome the limits of wildlife-habitat relationship studies. Journal of Wildlife Management **65**:613-623.

Morrison, M. L., B. G. Marcot, and R. W. Mannan. 1998. Wildlife habitat-relationships: concepts and applications. Second edition. University of Wisconsin Press, Madison, USA.

Mielke, P. W., Jr., and K. J. Berry. 2001. Permutation Methods: A Distance Function Approach. Springer-Verlag, New York. 352pp.

Neter, J., M. H. Kutner, C. J. Nachtsheim, and W. Wasserman. 1996. Applied linear statistical models. Irwin, Chicago, Illinois, USA. 1408pp.

O'Connor, R. J. 2002. The conceptual basis of species distribution modeling: Time for a paradigm shift? Pages 25-33 *in* J. M. Scott et al., editors. Predicting species occurrences: issues of accuracy and scale. Island Press, Covelo, California, USA.

Regan, H. M., M. Colyvan, and M. A. Burgman. 2002. A taxonomy and treatment of uncertainty for ecology and conservation biology. Ecological Applications **12**:618-628.

Rosenbaum, P. R. 1991. Discussing hidden bias in observational studies. Annals of Internal Medicine **115**:901-905.

Rosenbaum, P. R. 1995. Quantiles in nonrandom samples and observational studies. Journal of the American Statistical Association **90**:1424-1431.

Rosenbaum, P. R. 1999. Reduced sensitivity to hidden bias at upper quantiles in observational studies with dilated treatment effects. Biometrics **55**:560-564.

Rotenberry, J. T. 1986. Habitat relationships of shrubsteppe birds: even "good" models cannot predict the future. Pages 217-221 *in* J. Verner, M. L. Morrison, and C. j. Ralph, eds. Wildlife 2000: modeling habitat relationships of terrestrial vertebrates. Univ. Wisconsin Press, Madison

Scharf, F. S., F. Juanes, and M. Sutherland. 1998. Inferring ecological relationships from the edges of scatter diagrams: comparison of regression techniques. Ecology **79**:448-460.

Stauffer, D. F. 2002. Linking populations and habitats: Where have we been? Where are we going? Pages 53-61 *in* J. M. Scott et al., editors. Predicting species occurrences: issues of accuracy and scale. Island Press, Covelo, California, USA.

Terrell, J. W., B. S. Cade, J. Carpenter, and J. M. Thompson. 1996. Modeling stream fish habitat limitations from wedged-shaped patterns of variation in standing stock. Transactions of the American Fisheries Society **125**:104-117.

Terrell, J. W., and J. Carpenter. 1997. Selected habitat suitability index model evaluations. U. S. Department of Interior, Geological Survey, Information and

Technology Report USGS/BRD/ITR--1997--0005. 62pp.

Thomson, J. D., G. Weiblen, B. A. Thomson, S. Alfaro, and P. Legendre. 1996. Untangling multiple factors in spatial distributions: Lilies, gophers, and rocks. Ecology **77**:1698-1715.

Van Horne, B., and J. A. Wiens. 1991. Forest bird habitat suitability models and the development of general habitat models. U. S. Department of Interior, Fish and Wildlife Service, Fish and Wildlife Research 8. 31pp.

Vardeman, S. B. 1992. What about the other intervals? The American Statistician **46**: 193-197.

Wiens, J. A. 1989. The ecology of bird communities. Vol. 1. Cambridge Studies in Ecology, Cambridge University Press. 539pp.

Appendix 4

Appendix 4.1.

The $R^1(\tau)$ coefficient of determination was the proportionate reduction in the objective function minimized when passing from a constrained parameter quantile regression model to some unconstrained parameter model (Koenker and Machado 1999). My implementation of $R^1(\tau) = 1 - (SAF(\tau)/SAR(\tau))$ used

$SAR(\tau) = \min[\sum_{i=1}^{n} \rho_\tau(y_i - X_{0i}^{/}b_0)]$ for the reduced parameter model constrained to just a constant and used $SAF(\tau) = \min[\sum_{i=1}^{n} \rho_\tau(y_i - X_i^{/}b)]$ for the unconstrained full parameter model. This coefficient of determination was identical to the one used by Cade and Richards (1996) when $\tau = 0.50$.

The $\mathrm{AIC_c}(\tau) = -2\, l(\tau) + 2p(n/(n - p - 1))$, where $l(\tau)$ was the log-likelihood for the τth regression quantile and p was the number of parameters in the model (Hurvich and Tsai 1990). The likelihood used in the regression quantile $\mathrm{AIC_c}(\tau)$ assumed a double exponential distribution with density function $f_\sigma(e) = \tau(1 - \tau)\exp[-\rho_\tau(e)/\sigma]/\sigma$ and variance σ^2, where $\rho_\tau(e) = e(\tau - I(e < 0))$ was the check function used in minimizing the asymmetrically weighted sum of absolute deviations for regression quantiles (Koenker and Machado 1999). The log-likelihood $l(\tau) = n\ln(\tau(1 - \tau)) - n\ln\sigma - \sigma^{-1}[n\rho_\tau(e)]$ and $-2l(\tau)$ with $SAF(\tau)/n$ as an estimate of σ plugged in reduced to $-2n\ln(\tau(1 - \tau))$ +

$2n\ln(\mathrm{SAF}(\tau)/\mathrm{n}) + 2n$, where $SAF(\tau)$ was the weighted sum of absolute deviations minimized for the τth regression quantile estimate as above (Hurvich and Tsai 1990). In my implementation of $\mathrm{AIC}_c(\tau)$ to compare among models by quantile τ, I eliminated the terms $-2n\ln(\tau(1 - \tau)) + 2n$ because they were constants for any specified τ and, thus, cancelled when computing differences in $\mathrm{AIC}_c(\tau)$ between models by quantile [$\Delta\mathrm{AIC}_c(\tau)$].

Limited simulation work by Hurvich and Tsai (1990) and McQuarrie and Tsai (1998) indicated that model selection based on AIC_c for the 0.50 regression quantile was insensitive to occurrence of other error distributions than the double exponential assumed by the likelihood computations. Likelihoods for quantile regression for distributions other than the double exponential involve the multiplicative term $\sigma(\tau)/[\tau(1 - \tau)s(\tau)]$, where $s(\tau) = 1/f(F^{-1}(\tau))$ is the quantile density function (Koenker and Machado 1999). Since these terms would be constants in the likelihoods when comparing models using $\mathrm{AIC}_c(\tau)$ that assumed a common error distribution other than the double exponential, they would be irrelevant to the computed differences ($\Delta\mathrm{AIC}_c(\tau)$). The small sample, parameter penalty term in $\mathrm{AIC}_c(\tau)$, $2p(n/(n - p - 1))$, was based on normal distribution assumptions for least squares regression. Hurvich and Tsai (1990) and McQuarrie and Tsai (1998) found that more complex penalty terms suited for least absolute deviation regression and double exponential error distributions did not yield improved performance over the simpler term in AIC_c.

Appendix 4.2.

An example of computations for the quantile interval weights based on a modification of the method proposed by Koenker and Machado (1999) is provided for the 0.90 quantile for the model including bed elevation and bed elevation^2. The Hall and Sheather (1988) bandwidth rule assuming a normal distribution (for convenience) is $h(\tau) = n^{-1/3} z_\alpha^{2/3}[1.5\varphi^2(\Phi^{-1}(\tau))/2(\Phi^{-1}(\tau))^2 + 1)]^{1/3}$, where z_α satisfies $\Phi(z_\alpha) = 1 - \alpha/2$, Φ is the cdf and φ is the pdf of the standard normal distribution; and for the 0.90 quantile, α = 0.10, and n = 200 yielded a recommended bandwidth of $h(0.90) = 0.05264$. The estimates $b_0(\tau)$, $b_1(\tau)$, and $b_2(\tau)$ were obtained for all quantiles in the interval 0.90 ± $h(0.90) \in [0.84736, 0.95264]$. This interval contained 22 regression quantile estimates, and the average pairwise difference between them was 81.0003 for $b_0(\tau)$, 56.5343 for $b_1(\tau)$, and 9.98316 for $b_2(\tau)$. Plots of $b_1(\tau)$ and $b_2(\tau)$ by τ were examined to determine the sign of the rates of change to assign to the estimated difference coefficients. For this quantile the weights were $w(0.90) = (2 \times 0.05264)/(81.0031 - 56.5343 \times$ bed elevation $+ 9.9832 \times$ bed elevation^2). Plots of the weights as a function of bed elevation were examined to check for any negative weights; none occurred for $w(0.90)$. When negative weights were encountered a small constant was added to the denominator of the function to shift them all to positive values while preserving their relative value. The weights were then multiplied by *Macomona* >15 mm counts (y), bed elevation (X_1) and bed elevation^2 (X_2) to estimate the 0.90 quantile regression for the model $w(0.90)y = w(0.90) + w(0.90)X_1 + w(0.90)X_2$ and to compute confidence intervals based on inverting the quantile rankscore tests.